



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 114903

TO: Ruixiang Li
Location: REM-4D75
Art Unit: 1646
Thursday, February 26, 2004

Case Serial Number: 09742684

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Li,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

Thu Feb 26 13:51:32 2004

us-09-742-684a-15.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2004, 02:42:01 ; Search time 3101 Seconds
(without alignments)
2981.599 Million cell updates/sec

Title: US-09-742-684A-15
Perfect score: 2563
Sequence: 1 ctccaggaagaccagga.....aacacgttacaataagcc 2563

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2353733 seqs, 1803733377 residues

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 2556.4 | 99.7 | 2563 | 9 | US-09-742-684-1 |
| 2 | 1978.2 | 77.2 | 5286 | 14 | US-10-240-965-122 |
| 3 | 1803 | 74.2 | 2382 | 14 | US-10-101-510-140 |
| 4 | 678.2 | 26.5 | 2335 | 9 | US-09-742-684-3 |
| 5 | 403.6 | 15.7 | 452 | 9 | US-09-878-178-688 |
| 6 | 403.6 | 15.7 | 452 | 13 | US-10-046-935-688 |
| 7 | 403.6 | 15.7 | 452 | 14 | US-10-146-502-688 |
| 8 | 316.8 | 12.4 | 2687 | 13 | US-10-108-605-156 |
| 9 | 316.8 | 12.4 | 2687 | 13 | US-10-108-605-78 |
| 10 | 238.2 | 9.3 | 535 | 10 | US-09-918-995-30472 |
| 11 | 233 | 9.1 | 517 | 10 | US-09-918-995-32124 |
| 12 | 225.6 | 8.8 | 504 | 10 | US-09-918-995-24126 |
| 13 | 168.2 | 6.6 | 2090 | 9 | US-09-878-905-10 |
| 14 | 168.2 | 6.6 | 2090 | 14 | US-10-101-510-18 |
| 15 | 168.2 | 6.6 | 3206 | 14 | US-10-101-510-438 |

| | | | | | | |
|----|-------|-----|------|----|---------------------|---------------------|
| 16 | 168.2 | 6.6 | 5759 | 9 | US-09-954-456-2001 | Sequence 2001, App |
| 17 | 168.2 | 6.6 | 5821 | 13 | US-10-071-766-52 | Sequence 52, Appl |
| 18 | 160.8 | 6.3 | 1952 | 9 | US-09-903-068-17 | Sequence 17, Appl |
| 19 | 160.8 | 6.3 | 1952 | 9 | US-09-982-543A-7 | Sequence 7, Appl |
| 20 | 153.2 | 6.0 | 1813 | 9 | US-09-874-628-1 | Sequence 1, Appl |
| 21 | 151.6 | 5.9 | 2076 | 9 | US-09-874-628-3 | Sequence 3, Appl |
| 22 | 151.6 | 5.9 | 3003 | 15 | US-10-463-190-121 | Sequence 121, App |
| 23 | 151.6 | 5.9 | 3167 | 15 | US-10-463-190-119 | Sequence 119, App |
| 24 | 151.6 | 5.9 | 3167 | 15 | US-10-463-190-120 | Sequence 120, App |
| 25 | 148 | 5.8 | 2070 | 9 | US-09-903-068-13 | Sequence 13, Appl |
| 26 | 148 | 5.8 | 2402 | 10 | US-09-742-153-11 | Sequence 11, Appl |
| 27 | 147.8 | 5.8 | 1575 | 15 | US-10-463-190-117 | Sequence 117, App |
| 28 | 147.8 | 5.8 | 2032 | 12 | US-10-058-270A-5 | Sequence 5, Appl |
| 29 | 147.8 | 5.8 | 2032 | 13 | US-10-044-716-13 | GENERAL INFORMATION |
| 30 | 147.8 | 5.8 | 2032 | 14 | US-10-266-152A-39 | Sequence 39, Appl |
| 31 | 147.8 | 5.8 | 2032 | 14 | US-10-138-814-13 | Sequence 13, Appl |
| 32 | 147.8 | 5.8 | 2032 | 15 | US-10-295-027-67 | Sequence 67, Appl |
| 33 | 147.8 | 5.8 | 2032 | 15 | US-10-295-027-788 | Sequence 788, App |
| 34 | 147.8 | 5.8 | 2032 | 15 | US-10-295-027-836 | Sequence 836, App |
| 35 | 147.8 | 5.8 | 2032 | 15 | US-10-295-027-880 | Sequence 880, App |
| 36 | 147.8 | 5.8 | 2032 | 15 | US-10-173-999-63 | Sequence 63, Appl |
| 37 | 147.8 | 5.8 | 2032 | 15 | US-10-463-190-118 | Sequence 118, App |
| 38 | 147.8 | 5.8 | 2330 | 14 | US-10-198-846-10344 | Sequence 10344, A |
| 39 | 147.8 | 5.8 | 3722 | 14 | US-10-241-220-56 | Sequence 56, Appl |
| 40 | 147.2 | 5.7 | 1793 | 10 | US-09-069-228-1 | Sequence 1, Appl |
| 41 | 142.8 | 5.6 | 1602 | 9 | US-09-742-684-11 | Sequence 11, Appl |
| 42 | 142.8 | 5.6 | 1612 | 14 | US-10-169-051-3 | Sequence 3, Appl |
| 43 | 141.2 | 5.5 | 1612 | 14 | US-10-169-051-1 | Sequence 1, Appl |
| 44 | 141.2 | 5.5 | 1964 | 15 | US-10-463-190-123 | Sequence 123, App |
| 45 | 141.2 | 5.5 | 2063 | 15 | US-10-463-190-122 | Sequence 122, App |

ALIGNMENTS

RESULT 1
US-09-742-684-1
Sequence 1, Application US/09742684
Patent No. US20010039036A1
GENERAL INFORMATION:
APPLICANT: Mathews, Lawrence S.
Inventor: Tsuchida, Kunihiko
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF
RECEPTOR(S) OF THE ACTIVIN/TGF-BETA SUPERFAMILY
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSES: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/742,684
FILING DATE: 19-Dec-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,123
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/300,584
FILING DATE: 02-SEP-1994
APPLICATION NUMBER: US 07/880,220
FILING DATE: 08-MAY-1992
APPLICATION NUMBER: US 07/773,229
FILING DATE: 09-OCT-1991
APPLICATION NUMBER: US 07/698,709
FILING DATE: 10-MAY-1991

ATTORNEY/AGENT INFORMATION:

NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9927
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2563 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 71..1609
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-742-684-1

Query Match 99.7%; Score 2556.4; DB 9; Length 2563;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2557; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCGAGAGAGACCCAGGAACTGATATCTAGCGAGAACTTCTACGGCTTCTCGGCG 60
DB 1 CTCGAGAGAGAGACCCAGGAACTGATATCTAGCGAGAACTTCTACGGCTTCTCGGCG 60
QY 61 CCTCGGAGAAATGGAGAGCTGCTGCAAGTTGGCCGTCCTTCTTATCTCTTCTC 120
DB 61 CCTCGGAGAAATGGAGAGCTGCTGCAAGTTGGCCGTCCTTCTTATCTCTTCTC 120
QY 121 TTGAGGTCATATCTTGGCAGATCAGAACTCAGGAGTCTTCTTAAATGCTAATG 180
DB 121 TTGAGGTCATATCTTGGCAGATCAGAACTCAGGAGTCTTCTTAAATGCTAATG 180
QY 181 GGAAG 240
DB 181 GGAAG 240
QY 241 GCGACATTTGTTTCTACCTGGAAGATATTTCTGTTCCATTGAAATAGTGAAGAG 300
DB 241 GCGACATTTGTTTCTACCTGGAAGATATTTCTGTTCCATTGAAATAGTGAAGAG 300
QY 301 TTGTTGGCTGATATCACTGCTATGACAGAGCTGATGATGATGATGATGATGATG 360
DB 301 TTGTTGGCTGATATCACTGCTATGACAGAGCTGATGATGATGATGATGATGATG 360
QY 361 CCCGAGAGTACTTGTGCTGAGAGGCAATATGATGATGATGATGATGATGATGATG 420
DB 361 CCCGAGAGTACTTGTGCTGAGAGGCAATATGATGATGATGATGATGATGATGATG 420
QY 421 TCCGAGATGAG 480
DB 421 TCCGAGATGAG 480
QY 481 CAACATTTCTGCTGATTTCTGTTGATCACTAATGTTAATGAGAGATTTGCTATTG 540
DB 481 CAACATTTCTGCTGATTTCTGTTGATCACTAATGTTAATGAGAGATTTGCTATTG 540
QY 541 ATTTTGGGTGACAGACATCAACAGATGGCTTACCTCTGTAATGAGAGAGAGAGAG 600
DB 541 ATTTTGGGTGACAGACATCAACAGATGGCTTACCTCTGTAATGAGAGAGAGAGAG 600
QY 601 CCCGAG 660
DB 601 CCCGAG 660
QY 661 GAAAG 720
DB 661 GAAAG 720
QY 721 TGTCAAAATATTTTCCATATACAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGATG 780

DB 721 TGTCAAAATATTTTCCATATACAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGATG 780
QY 781 TCTACCTGAGATGAG 840
DB 781 TCTACCTGAGATGAG 840
QY 841 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 901 CTTTCTTAAAGGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 CTTTCTTAAAGGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 TAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 TAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 AATCTCTCAG 1080
DB 1021 AATCTCTCAG 1080
QY 1081 CATTTGCTGACTTTGGGCTTAAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 CATTTGCTGACTTTGGGCTTAAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 TGGGAG 1200
DB 1141 TGGGAG 1200
QY 1201 CCAAG 1260
DB 1201 CCAAG 1260
QY 1261 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 AGAATTTGGGAG 1380
DB 1321 AGAATTTGGGAG 1380
QY 1381 GAGGCTGTTTAAAGAGATTTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1381 GAGGCTGTTTAAAGAGATTTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 GATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1441 GATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1501 AAGAAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB 1501 AAGAAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
QY 1561 GGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 GGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 CCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
DB 1621 CCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
QY 1681 TTAGTGAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1681 TTAGTGAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
QY 1741 CTTGAG 1800
DB 1741 CTTGAG 1800
QY 1801 CAGATATGAG 1860

Db 1801 CAGATATGAGAGAGCTTAAGGAAAAAGCTGAACTGTAAAGAACTTCTGAAAAATGTA 1860
 QY 1861 CTCGAGAGATGAGGCGCTCCCAATCAAGAGATCTTTTGAACCTGGCTAATCAAGTATTT 1920
 Db 1861 CTCGAGAGATGAGGCGCTCCCAATCAAGAGATCTTTTGAACCTGGCTAATCAAGTATTT 1920
 QY 1921 GCAAACTGACATCAGATTTCTTAATGTCTGTGAGAGACATTAATTCCTTAATGAAGT 1980
 Db 1921 GCAAACTGACATCAGATTTCTTAATGTCTGTGAGAGACATTAATTCCTTAATGAAGT 1980
 QY 1981 ACTGCTATTTTCTTAATGAAAAAATTTTCAATTTCAATTTTAAAAAGGTTAACTTTT 2040
 Db 1981 ACTGCTATTTTCTTAATGAAAAAATTTTCAATTTCAATTTTAAAAAGGTTAACTTTT 2040
 QY 2041 ATTGCAATTTGCTGTGTTTCTAATGAATGATGATGATGATGATGATGATGATGATGATG 2100
 Db 2041 ATTGCAATTTGCTGTGTTTCTAATGAATGATGATGATGATGATGATGATGATGATG 2100
 QY 2101 AATGTAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2160
 Db 2101 AATGTAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2160
 QY 2161 CTCGAGAGATTTACTTTTAACTCCCTCAACAGAGTAACTCCAGTTCACGGTTGTTAA 2220
 Db 2161 CTCGAGAGATTTACTTTTAACTCCCTCAACAGAGTAACTCCAGTTCACGGTTGTTAA 2220
 QY 2221 TTATTAATTTGAAAAACACTTAACAGATTTGAATTAATCACTGATTTTAAATCAAGT 2280
 Db 2221 TTATTAATTTGAAAAACACTTAACAGATTTGAATTAATCACTGATTTTAAATCAAGT 2280
 QY 2281 TAATTAACAATTTCACTGTGTATTTAAGAAAAAATGTTAAGTAACTAGTAACTAGTAACT 2340
 Db 2281 TAATTAACAATTTCACTGTGTATTTAAGAAAAAATGTTAAGTAACTAGTAACTAGTAACT 2340
 QY 2341 TAAATGCTATTTGTAAGAGAGCTGTTTAACTTTCTTCTGCTGCTGCTGCTGCTGCTGCT 2400
 Db 2341 TAAATGCTATTTGTAAGAGAGCTGTTTAACTTTCTTCTGCTGCTGCTGCTGCTGCTGCT 2400
 QY 2401 AAACAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 Db 2401 AAACAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 QY 2461 ATCAAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 Db 2461 ATCAAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 QY 2521 CAAGCCAGTAGAAGAAAGCTTAAACACGCTTTTCAAAATAGCC 2563
 Db 2521 CAAGCCAGTAGAAGAAAGCTTAAACACGCTTTTCAAAATAGCC 2563

RESULT 2
 us-10-240-965-122
 ; Sequence 122, Application US/10240965
 ; Publication No. US20030165924A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: SHIFFMAN, DOV
 ; APPLICANT: SOMOGYI, Roland
 ; APPLICANT: LAMN, Richard M.
 ; APPLICANT: SEILHAMER, Jeffrey J.
 ; APPLICANT: PORTER, Gordon J.
 ; APPLICANT: WIKITA, Thomas
 ; APPLICANT: TAI, Julie
 ; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
 ; FILE REFERENCE: PA-0025 PCT
 ; CURRENT APPLICATION NUMBER: US/10/240,965
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: 60/195,106
 ; NUMBER OF SEQ ID NOS: 276
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 122

; LENGTH: 5286
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030165924A1 014704.3
 ; NAME/KEY: unsure
 ; LOCATION: 4789-4822
 ; OTHER INFORMATION: a, t, c, g, or other
 ; US-10-240-965-122
 Query Match 77.2%; Score 1978.2; DB 14; Length 5286;
 Best Local Similarity 91.0%; Pred. No. 0;
 Matches 2222; Conservative 2; Mismatches 181; Indels 37; Gaps 10;
 QY 1 CTCGAGAGAAACCCAGGAGCTGATCTAGAGAGAACTCTACGGCTTCCGGAGC 60
 Db 103 CTCGAGAGAAACCCAGGAGCTGATCTAGAGAGAACTCTCTCGAATTCGCCGCG 162
 QY 61 CCTGGGAAATGGAGCTGCTGCAAGTTGGCGCTTCTTATCTCTGCTC 120
 Db 163 CCTGGGAAATGGAGCTGCTGCAAGTTGGCGCTTCTTATCTCTGCTC 222
 QY 121 TTCAAGTCTATCTTGGCAGATCAGAACTCAGAGATGCTTTTCTTAATGCTAATG 180
 Db 223 TTCAAGTCTATCTTGGCAGATCAGAACTCAGAGATGCTTTTCTTAATGCTAATG 282
 QY 223 TTCAAGTCTATCTTGGCAGATCAGAACTCAGAGATGCTTTTCTTAATGCTAATG 282
 Db 181 GAAAAAGACAGAACCAACCAAGCTGTGTGAACCTTGTATGTATGAATTAACG 240
 QY 283 GAAAAAGACAGAACCAACCAAGCTGTGTGAACCTTGTATGTATGAATTAACG 342
 Db 241 GCGACATTTTGTGCTACCTGGAAGAAATTTCTGCTTCAATGAAATGTAAGGAAG 299
 QY 343 GCGGATTTGTTGCTTCACTGGAAGAAATTTCTGCTTCAATGAAATGTAAGGAAG 402
 Db 300 GTTGTGCTGATATATCACTGCTATGACAGAGCTGATGTTGTTGAAAAAAGACA 359
 QY 403 GTTGTGCTGATATATCACTGCTATGACAGAGCTGATGTTGTTGAAAAAAGACA 462
 Db 360 GCGCTGAAGTGAATTTTGTGCTGTGAAGGCAATGTGTAAATGAAGTCTCTTATT 419
 QY 463 GCGCTGAAGTGAATTTTGTGCTGTGAAGGCAATGTGTAAATGAAGTCTCTTATT 522
 Db 420 TTCCGAGATGGAAGTCAACAGCCCACTTCAATCTGTTTACACCAAGCCACCTATT 479
 QY 523 TTCCGAGATGGAAGTCAACAGCCCACTTCAATCTGTTTACACCAAGCCACCTATT 582
 Db 480 ACAATCTGCTGTATTTCTGTTGATGCACTAATGTTAATGGAAGAAATGCTATTGTG 539
 QY 583 ACAATCTGCTGTATTTCTGTTGATGCACTAATGTTAATGGAAGAAATGCTATTGTG 642
 Db 540 CATTTGGGTGTAAGACATCAACAGATGGCTTACCTCTGTAATTTCTTCTACTCAAG 599
 QY 643 CATTTGGGTGTAAGACATCAACAGATGGCTTACCTCTGTAATTTCTTCTACTCAAG 702
 Db 600 ACCCAGACCAACCCCACTTCCCATTAAGGTTGAAGCCATTGCAAGCTGTTAAG 659
 QY 703 ACCCAGACCAACCCCACTTCCCATTAAGGTTGAAGCCATTGCAAGCTGTTAAG 762
 Db 660 TGAAGCAAGGGAGATTTGTTGTGTCTGAAGAGCCAGTGTCTCAATGAATGTGG 719
 QY 763 TGAAGCAAGGGAGATTTGTTGTGTCTGAAGAGCCAGTGTCTCAATGAATGTGG 822
 Db 720 CTGTCAAAATATTTTCAATACAGACAAAGTCCCTGCAAGATGATGATGATGATG 779
 QY 823 CTGTCAAAATATTTTCAATACAGACAAAGTCCCTGCAAGATGATGATGATGATG 882
 Db 780 GTCTACCTGGAATGAAGCATGAGAACTATCACTGATTTGTTGCAAGAAAGAGCA 839
 QY 883 GTTGTGCTGGAATGAAGCATGAGAACTATTAAGTTCATTTGTTGCAAGAAAGAGCA 942
 Db 840 CCAGTGTGATGTGACCTGTGTATCAACAGATTTCAATGAAAAAGGCTCAGTGTAG 899

| | | | |
|----|------|--|------|
| Db | 284 | GGAAAAAGACAGAACCAATCAACTGGTGTGGAACCGGTGTATGGTACAAAGATTAAG | 343 |
| QY | 241 | GCACATTGTTTGTCTACCTGGAAGATATTTCTGGTTCATTGAATATGTGAAGCAAG | 300 |
| Db | 344 | GCGGCAATTGTTTGTCTACCTGGAAGATATTTCTGGTTCATTGAATATGTGAAGCAAG | 403 |
| QY | 301 | TTGTTGGCTGAGATGATATCACTGCTATGACAGACTGATTTGTGNGAAAAAAGACAG | 360 |
| Db | 404 | TTGTTGGCTGAGATGATATCACTGCTATGACAGACTGATTTGTGNGAAAAAAGACAG | 463 |
| QY | 361 | CCCTGAAGTACTTTTGTGTGCTGTAGGGGCAATATGTGTAATGAAGTTCTCTAATT | 420 |
| Db | 464 | CCCTGAAGTACTTTTGTGTGCTGTAGGGGCAATATGTGTAATGAAGTTCTCTAATT | 523 |
| QY | 421 | TCCGGAGATGGAAGTACACAGAGCCACTTCAAATCCCTGTATCACACGGAAGCCACTATTA | 480 |
| Db | 524 | TCCGGAGATGGAAGTACACAGAGCCACTTCAAATCCCTGTATCACACGGAAGCCACTATTA | 583 |
| QY | 481 | CAACATTCGCTGTATTCCTTGGTACCACTATATGTAATTCAGAGAAATGTCAATTGTGC | 540 |
| Db | 584 | CAACATTCGCTGTATTCCTTGGTACCACTATATGTAATTCAGAGAAATGTCAATTGTGC | 643 |
| QY | 541 | ATTTTGGGTATCAGACATCACAAGATGGCCTACCTCCTGTACTTGTTCCTACTCAAG | 600 |
| Db | 644 | ATTTTGGGTATCAGACATCACAAGATGGCCTACCTCCTGTACTTGTTCCTACTCAAG | 703 |
| QY | 601 | CCGAGACCAACCCCACTTCCTCCCATTACTAGGGTTGAAGCCATTGACACTGTTAAGT | 660 |
| Db | 704 | CCGAGACCAACCCCACTTCCTCCCATTACTAGGGTTGAAGCCATTGACACTGTTAAGT | 763 |
| QY | 661 | GAAAGCAAGGGGAAGATTTGGTGTGTCTGGAAGCCCAATTGCTCAATGATATGTGC | 720 |
| Db | 764 | GAAAGCAAGGGGAAGATTTGGTGTGTCTGGAAGCCCAATTGCTCAATGATATGTGC | 823 |
| QY | 721 | TGTCAAAATATTTCCATATACAGACAACACATCCCGGAGAAATGAATAGAAGTCTATAG | 780 |
| Db | 824 | TGTCAAAATATTTCCATATACAGACAACACATCCCGGAGAAATGAATAGAAGTCTATAG | 883 |
| QY | 781 | TCTACCTGGAATGAAGCATGAAGACATACATACTACAGTTATGCTGACAGAAAAAGAGCAC | 840 |
| Db | 884 | TTTGCTGTGAATGAAGCATGAAGACATATTAAGTTCAATGCTGACAGAAAAAGAGCAC | 943 |
| QY | 841 | CAGTGTGATGTGACCTGTGGCTAATCACAGCAATTCATGAAAAAGGCTCACTGTGAGA | 900 |
| Db | 944 | CAGTGTGATGTGACCTGTGGCTAATCACAGCAATTCATGAAAAAGGCTCACTGTGAGA | 1003 |
| QY | 901 | CTTTCTTAAAGGCTAATGTGCTCTTGGAAATCARTTTGTCAATTTGCAGAAACCAATGCG | 960 |
| Db | 1004 | CTTTCTTAAAGGCTAATGTGCTCTTGGAAATCARTTTGTCAATTTGCAGAAACCAATGCG | 1063 |
| QY | 961 | TGAGAGATTGSCATATTTACATGAGGATATACCTGGCTTAAAGATGGGCCAAGCCTGC | 1020 |
| Db | 1064 | TGAGAGATTGSCATATTTACATGAGGATATACCTGGCTTAAAGATGGGCCAAGCCTGC | 1123 |
| QY | 1021 | AATCTCTCAGGGACATCMAAAATTAATAAATGTGCTGTGAAAAAACAATCTGACAGTTG | 1080 |
| Db | 1124 | CATATCTCAGGGACATCMAAAATTAATAAATGTGCTGTGAAAAAACAATCTGACAGTTG | 1183 |
| QY | 1081 | CATTGCTGAATTTGGGTGGCTTAAAGTTGAGAGCTGGCAAGTCTGACAGTACACCCA | 1140 |
| Db | 1184 | CATTGCTGAATTTGGGTGGCTTAAAGTTGAGAGCTGGCAAGTCTGACAGTACACCCA | 1243 |
| QY | 1244 | TGAGACAGTTGGTATCCGGAGGTATCATGGCTCAAGAGGTATTAAGAGGTGTCTAATACTT | 1303 |
| Db | 1201 | CCAAAGGAGACGCAATTTCTGAGAGATGATATGTACGCATGGGATTTGTCTTATGGCAAT | 1260 |
| QY | 1304 | CCAAAGGAGATGCAATTTTGTAGAGATGATATGTATGTCCATGGGATTTGTCTTATGGCACT | 1363 |
| Db | 1261 | GGCTTCTCGTTGCACTGCTGAGATGAGACCCCTAATATGATACATTTTACCATTGTAGGA | 1320 |

| | | | |
|----|------|---|------|
| Db | 1364 | GGCTTCGCGTGAACGTGCTCAAGATGGAACCTGTAAAGATGAATACATGTTGCCATTTGAGGA | 1423 |
| Qy | 1321 | AGAAATTGGCCAGCATTCATCTCTTGAAGATATGACGAAGATTGTGTGCATAAAAAA | 1380 |
| Db | 1424 | GGAAATTGGCCAGCATTCATCTCTTGAAGATATGACGAAGATTGTGTGCATAAAAAA | 1483 |
| Qy | 1381 | GAGGCGCTTTTAAAGATTAATTTGGGCAAGAACATGCAAGAAATGGCAATGCTGTGAAC | 1440 |
| Db | 1484 | GAGGCGCTTTTAAAGATTAATTTGGGCAAGAACATGCTGAATGGCAATGCTGTGAAC | 1543 |
| Qy | 1441 | GATGAAGAAATGTTGGGATCATGATGCAAGACGAGTTATCAAGTGGATGTGAGTGA | 1500 |
| Db | 1544 | CATTGAAGAAATGTTGGGATCAAGACGAGAGCAAGGTTATCAAGTGGATGTGAGTGA | 1603 |
| Qy | 1501 | AAGAAATTAATCAATGCAAGCACTAAATCAAAATATCAATTAACAAGACATTTGAACAT | 1560 |
| Db | 1604 | AAGAAATTAATCAATGCAAGCACTAAATCAAAATATTAACAAGACATTTGAACAT | 1663 |
| Qy | 1561 | GGTCACAATGATGCAAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGTGGCA | 1620 |
| Db | 1664 | GGTCACAATGATGCAAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGTGGCG | 1723 |
| Qy | 1621 | CCGCTCTGTACACATGAGGACTGGGACTCTGAACTGAGAGCTGCTAACTAAGAAAGTGC | 1680 |
| Db | 1724 | CCATCTGTGTGACACTAAGAAATGGGACTCTGAACTGAGAGCTGCTAACTAAGAAAGTGC | 1783 |
| Qy | 1681 | TT--AGTTGAATTTCTGTGGAATGATGATGAGATGCGCTCCAGACATGTACGCAAGCAGC | 1738 |
| Db | 1784 | TTACAGTTATTTCTGTGGAATGATGATGAGATGCTCTTGGAAATGTTAAGAAAGAG | 1843 |
| Qy | 1739 | CCCTTGTGGAAGCATGATCTGGGAAATGGATCTGGGAAACTTAATCAATCGTCTGCAG | 1798 |
| Db | 1844 | ACCTTTGTGGAAG-------AGTTGCTCTGGGAGACTTACGTGCATTTGCCAGAG | 1893 |
| Qy | 1799 | CACAGATATGAG--AGAGCTTAAGGAAAGACCTGCACAACTGTAAAGAACTTCTG-- | 1852 |
| Db | 1894 | CACAGATATGAGGACATGAGATGTAAAGAAAGACCTGCACAACTGTAAAGAACTTCTG | 1953 |
| Qy | 1853 | --AAATGTACTGAAGATGTGGCCTCTCCAAATCAAGATCTTTTGGACCTGGCTAA | 1910 |
| Db | 1954 | AAAAGTGTACTGAAGATGTGGCCTCTCCAAATCAAGATCTTTTGGACCTGGCTAA | 2013 |
| Qy | 1911 | TCAAGTATTTGCAAAATGTGACATCAAGATTTCTTATGTGCTGAGAGACACTAATTCCT | 1970 |
| Db | 2014 | TGAGGTGTTG--AAAATGTGACATCAAGATTTCTTATGTGCTGAGAGACACTAATTCCT | 2072 |
| Qy | 1971 | TAAATGAACATACGTCTAATTTTAAATGAAAACTTTCAATTCAGATTTTAAAAAGG | 2030 |
| Db | 2073 | TAAATGAACATACGTCTAATTTTAAATGAAAACTTTCAATTCAGATTTTAAAAAGG | 2133 |
| Qy | 2031 | GTAAC--TTTATTTGCAATTTGCTGTGTTCTATAATGACTATTTGATGCGACAT | 2087 |
| Db | 2133 | GTAACGTTTATTTGCAATTTGCTGTGTTCTATAATGACTATTTGATGCGACAT | 2192 |
| Qy | 2088 | GACACAGCTTGTGAATGTAGTGTGCTGCTGTTCTGTGACAT--AGTCATCAAGTGG | 2145 |
| Db | 2193 | GACACAGCTTGTGAATGTAGTGTGCTGCTGTTCTGTGACATCAAGTGAAGTGG | 2252 |
| Qy | 2146 | GGTACAGTAAAGAGGCTTCCAGCATTAATTTAACTTCCCTCAACAAAGTATACCTGAGT | 2200 |
| Db | 2253 | GGTACAGTAAAGAGGCTTCCAGCATTAATTTAACTTCCCTCAACAAAGTATACCTGAGT | 2312 |
| Qy | 2206 | TCCACGGTGTGTTAATTAATTAATGAAAAACATCAAGCAATTTGAATAA | 2255 |
| Db | 2313 | TCCACGGTGTGTTAATTAATTAATGAAAAACATCAAGCAATTTGAATAA | 2362 |

```

Vale, Wylie W.
Teuchida, Kunihito
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF
RECEPTOR(S) OF THE ACTIVIN/TGF-BETA SUPERFAMILY
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/742,684
FILING DATE: 19-Dec-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,123
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/300,584
FILING DATE: 02-SEP-1994
APPLICATION NUMBER: US 07/880,220
FILING DATE: 08-MAY-1992
APPLICATION NUMBER: US 07/773,229
FILING DATE: 09-OCT-1991
APPLICATION NUMBER: US 07/658,709
FILING DATE: 10-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9927
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
SEQUENCE CHARACTERISTICS:
LENGTH: 2335 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: XACTR
FEATURE:
NAME/KEY: CDS
LOCATION: 468..1997
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-742-684-3
Query Match 26.5%; Score 678.2; DB 9; Length 2335;
Best Local Similarity 65.6%; Pred. No. 6,1e-169;
Matches 1037; Conservative 2; Mismatches 526; Indels 15; Gaps 3
62 CTCGGGAAAATGGGAGCTGCTGCACAAAGTTGGGGCTTCCGCGTCTTTCTTATCTCTGCTCT 121
459 CCCAGAGAGATGGGGGCGTGTAGGCGTGCACCTTTCTTACTCTTCTTGGAACTTTCGCGC 518
122 TCAGGTCATATCTTGGCAGATCAGAACTCAGAGTGTCTTTCTTTATGCTAATGG 181
519 GAGGCTCAGACAGATGAAGTGAAGACAAAGAGTGATCTATTACATGCCAATCG 578
182 GAAAAGACAGAACCAACAGACTGTGTGTTGAACCTTGTGTATGTGATTAAGATAACGG 241
579 GAACGTGAAGAACCAACCAAGTGGGTGGAAAGCTGCGAAGGGGAAAAGAGCAACAGCGA 638
242 CGACATTTGTTTGCTACTCTGGAAGAAATTTTCTGTGTTCCATTGAAATATGGAAGCAAGT 301
639 CTCACATGTTAGCCGCTTTGAGAGAACCAATTGGGCTTATGAGCTGTGAAAAGAAAAGGA 698

```

| | | | |
|----|------|---|------|
| QY | 302 | TGTTGGCTGGAGATGATATCAACTGCTCTATGACAGGACTGATTTGTGTGAAAAAAGACAC | 361 |
| Db | 699 | TGCTGCTGGATGACTCTCAACTCTTTATGACAGACAGGATGATTTGCCAAGGAAGAAAC | 758 |
| QY | 362 | CTGTAAAGTACTCTTTGTGTGCTGTGAGGGCAATGTGTAAAGAAAGTCTCTTAATTTT | 421 |
| Db | 759 | CCCCAGTCTTTTTCGTGCTGTGAGGAGAACTACTGCAACAAAGAAATTAATTAATCT | 818 |
| QY | 422 | CCGGAGATGAACTCAACAGGCCACTTCAATCTCTGTTACACCGAAGCCACTATAC | 481 |
| Db | 819 | CTTGAAGTCGAAACATTTGATCCGAGACCCCA-----GCCCTCAGCCCTCCGTA | 869 |
| QY | 482 | AACATTCCTGTAATTCCTTGGTAACCACTATGTTAATTTGACGAATGTCAATTTGTGA | 541 |
| Db | 870 | AACATTCATATCTATTCCTGCTTCCAAATGTGTGCTTTTCATGCAATTCCTCTGGG | 929 |
| QY | 542 | TTTTGGGTATACAGACATCAACAGATGACCTCCTCTGATCTTGTTCCTACTCAAC | 601 |
| Db | 930 | TTCTGATATGACCTCATCGAAGGCCCTCTAAGCGGCAATGTA---GAGTCAATGAGAC | 986 |
| QY | 602 | CCAGAACCAACCCCACTTCCCATTACTAGGGTTGAAGCCATTGACGCTGTTAAGT | 661 |
| Db | 987 | CCCGGTGCCCCCTCATCTCCTGCTGAGGCTGAAGCCCTGCAAGTTGCTGAGATA | 1046 |
| QY | 662 | AAAGCAAGGGAAAGATTTGGTTGTGCTGAAAGCCAGTTGCTCAATGAATATGAGCT | 721 |
| Db | 1047 | AAGGCCCGAGCGGTTTCGGTTGGGTCTGAAAGCTGCTGCTGATGAATATGTGCA | 1106 |
| QY | 722 | GTCAAAATATTTCCATATACAGGACAAACAGTCTGCGAATGAAATATGAAGTCTAGT | 781 |
| Db | 1107 | GTGAATATCTTCCCGTGCAGAGATTAAGCATGTGCGCACTGTGAAGAAAGATCTTACC | 1166 |
| QY | 782 | CTACCTGSAATGAAGCATGAGACATACTACATTCATTTGCTGACAGAAAGAGCACC | 841 |
| Db | 1167 | ACGCCCGGCAATGAACATGAACAACTATGTGAGTTCATTGCGGCTGAGAAAGGGGAGC | 1226 |
| QY | 842 | AGTGTGATGTGACCTGTGGCTATATCAAGCATTTTCAAGAAAGGGCTCACTGTGAC | 901 |
| Db | 1227 | AACCTGAGATGAGCTGTGGCTCATCTGCTATTTCAATGAAGGGTCTCTGACGAC | 1286 |
| QY | 902 | TTTCTTAAGGCTATGTGCTCTTTGATATCATCTTGTCAATTCAGAAACCATGCT | 961 |
| Db | 1287 | TACCTGAAGGAACTTGTGTAGCTGSAATGACTGTGTCACTAACAGAAACATGCT | 1346 |
| QY | 962 | AGAGATTTGCAATTTACATGAGGATATACCTGTCTTAAA--GATGGCAACAGCT | 1018 |
| Db | 1347 | CGTGGGTGGCTCTTACATGAAGATGTGCCCGCTGTAAAGGTAAAGGGCAACAACT | 1406 |
| QY | 1019 | GCAATCTCTACAGGACATCAAAAGTAAAGTGTGCTGTGAAAAACAATCTGACACT | 1078 |
| Db | 1407 | GCAATGCTCACAGAGATTTTAAAGTAAAGATGTATCTTAAAGAAACGACTGACTCG | 1466 |
| QY | 1079 | TGCATTCCTGACTTTGGTGTGGCTTAAAGTTCAAGCTGTGCAAGCTGTGAAGTGAAC | 1138 |
| Db | 1467 | ATATTACAGACTTCGGGCTGCGCTGACGATTTGAGCTGTGAAAACCTCCGGGAGATCA | 1522 |
| QY | 1139 | CATGGGAGGTTGTGTAACCGGAGGATATATGCTCCAGAGGTTGTGAGGCTGCTATAAC | 1198 |
| Db | 1527 | CACGGGAGGTTGTGACCAAGAGATATATGCTCTCGAGGTTCTTAAGGGAGCAATTAAC | 1586 |
| QY | 1199 | TTCCAAAGGACGATTTCTGAGGATGATATGTACGCCATGGGATTAATGCTTATGGAA | 1256 |
| Db | 1587 | TTTCAGAGAGATTCCTTTCTCAGGATATGATATGATGCCATGGGACTGTATCTGGGAA | 1646 |
| QY | 1259 | TTGGCTTCTGTTGACATGCTGCAATGGAACCCGTATGATGATGACATGTTACATTTGAG | 1316 |
| Db | 1647 | ATAGTATCCCATATGACAGACAGCAATGGGCCATGATGAGATCTGCTCCCATTCGAA | 1706 |
| QY | 1319 | GAAGAAATTTGGCCAGCATCATCTGTTGAAGATATCAGAAATGTTGTGTGATPAAAAA | 1376 |
| Db | 1707 | GAAGAAATTTGGCCAGCATCTTCTTCCATAGGATCTGCAAGAAATGTTGTGTGATPAAAAA | 1766 |

2Y 1379 AAGAGGCTGTTTAAAGATATTGGCAGAAACATGACAGGAATGCAATGCTCTGTGAA 1438
Db 1767 ATACGCCCTGTATTCAAGACACCTGGTGAAACACCCCTGGTCTGGCCCAACCTGTGGTC 1826
2Y 1439 ACATAGAAGAAATGTTGGATCATGATCAGAAAGCCAGGTATACAGTGTGATGTAGGT 1498
Db 1827 ACCAATGAAGAAATGCTGGACCATGATGCGAAGCAAGGCTTTCGCGAGGCTCGTAGAG 1886
2Y 1499 GAAAGAAATTAATCAGATGCAAGACATAACAAATATCATTTACTACAGAGGACATTTGAACA 1558
Db 1887 GAGCGTATTTCCCAATCCGTAATCAGTGAAGCGGCACTACCTCGGAGCTCCCTGTATCC 1946
2Y 1559 GTGGTCACAATGGTGACAAATGTTGACTTTCCTCCCAAGAAATCTAGTCTATGATGGTGG 1618
Db 1947 ATTGTTACATCTGTCAACAAATGTGGACTTCCCGCCCAAGAGTCCAGTATCTGAGGTTTC 2006
2Y 1619 CACCGTCTGTACACTGAG 1638
Db 2007 TTTGGTCTTTCAGACTCAG 2026

RESULT 5

US-09-878-178-688
; Sequence 688, Application US/09878178
; Patent No. US2002017552A1

GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.527

; CURRENT APPLICATION NUMBER: US/09/878,178

; CURRENT FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 2237

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 688

; LENGTH: 452

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-878-178-688

Query Match 15.7%; Score 403.6; DB 9; Length 452;
Best Local Similarity 93.1%; Pred. No. 1.8e-96;
Matches 421; Conservative 1; Mismatches 30; Indels 0; Gaps 0;

QY 583 ACTTGTTCTCTACTCAAGACCCAGGACCCACCCCTTCCCATTTACTAGGGTTGAAGCC 642
Db 1 ACTTGTTCTCTACTCAAGACCCAGGACCCACCCCTTCCCATTTACTAGGGTTGAAGCC 60
QY 643 ATTGCAGCTGTTAGAAAGTGAAGCAAGGGAAGATTTGGTTGTTCTGGAAGCCCAAGTT 702
Db 61 ACTGCAGTTATTAGAAGTGAAGCAAGGGAAGATTTGGTTGTTCTGGAAGCCCAAGTT 120
QY 703 GCTCAATGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGCAGAA 762
Db 121 GCTTAACGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGCAGAA 180
QY 763 TGAATATGAAGTCTATAGTCTACTGGAATGAAGCAATTTGGTTGTTCTGGAAGCCCAAGTT 822
Db 181 TGAATACGAAGTCTACAGTTTGGCTGGAATGAAGCAATTTGGTTGTTCTGGAAGCCCAAGTT 240
QY 823 TGCAGAGAAAGAGGACCAAGTGTGGATGTGGACCTGTGGCTTAATCAAGCAATTTTCATGA 882
Db 241 TGCAGAAAGAGGACCAAGTGTGGATGTGGATGTGGATGTGGCTGATCAAGCAATTTTCATGA 300
QY 883 AAAGGGCTCACTGTCAAGCTTTCTTAAGGCTTAATGTGGTCTCTTGGAAATCACTTTGTCA 942
Db 301 AAAGGGTTCACATCAAGCTTTCTTAAGGCTTAATGTGGTCTCTTGGAAATCACTTTGTCA 360
QY 943 TATTGAGAAACCAATGGCTAGAGGATTTGGCAATTTTACATGAGGATATACCTGGCTTAAA 1002
Db 361 TATTGAGAAACCAATGGCTAGAGGATTTGGCAATTTTACATGAGGATATACCTGGCTTAAA 420

QY 1003 AGATGGCCACAAGCCTGCAATCTCTCACAGGG 1034
Db 421 AGATGGCCACAACCTGCCATATCTCACAGGG 452

RESULT 6

US-10-046-935-688
; Sequence 688, Application US/10046935
; Publication No. US20020156011A1

GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.527C1

; CURRENT APPLICATION NUMBER: US/10/046,935

; CURRENT FILING DATE: 2002-01-15

; NUMBER OF SEQ ID NOS: 2239

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 688

; LENGTH: 452

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-046-935-688

Query Match 15.7%; Score 403.6; DB 13; Length 452;
Best Local Similarity 93.1%; Pred. No. 1.8e-96;
Matches 421; Conservative 1; Mismatches 30; Indels 0; Gaps 0;

QY 583 ACTTGTTCTCTACTCAAGACCCAGGACCCACCCCTTCCCATTTACTAGGGTTGAAGCC 642
Db 1 ACTTGTTCTCTACTCAAGACCCAGGACCCACCCCTTCCCATTTACTAGGGTTGAAGCC 60
QY 643 ATTGCAGCTGTTAGAAGTGAAGCAAGGGAAGATTTGGTTGTTCTGGAAGCCCAAGTT 702
Db 61 ACTGCAGTTATTAGAAGTGAAGCAAGGGAAGATTTGGTTGTTCTGGAAGCCCAAGTT 120
QY 703 GCTCAATGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGCAGAA 762
Db 121 GCTTAACGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGCAGAA 180
QY 763 TGAATATGAAGTCTATAGTCTACTGGAATGAAGCAATTTGGTTGTTCTGGAAGCCCAAGTT 822
Db 181 TGAATACGAAGTCTACAGTTTGGCTGGAATGAAGCAATTTGGTTGTTCTGGAAGCCCAAGTT 240
QY 823 TGCAGAGAAAGAGGACCAAGTGTGGATGTGGACCTGTGGCTTAATCAAGCAATTTTCATGA 882
Db 241 TGCAGAAAGAGGACCAAGTGTGGATGTGGATGTGGATGTGGCTGATCAAGCAATTTTCATGA 300
QY 883 AAAGGGCTCACTGTCAAGCTTTCTTAAGGCTTAATGTGGTCTCTTGGAAATCACTTTGTCA 942
Db 301 AAAGGGTTCACATCAAGCTTTCTTAAGGCTTAATGTGGTCTCTTGGAAATCACTTTGTCA 360
QY 943 TATTGAGAAACCAATGGCTAGAGGATTTGGCAATTTTACATGAGGATATACCTGGCTTAAA 1002
Db 361 TATTGAGAAACCAATGGCTAGAGGATTTGGCAATTTTACATGAGGATATACCTGGCTTAAA 420
QY 1003 AGATGGCCACAAGCCTGCAATCTCTCACAGGG 1034
Db 421 AGATGGCCACAACCTGCCATATCTCACAGGG 452

RESULT 7

US-10-146-502-688
; Sequence 688, Application US/10146502
; Publication No. US20030069180A1

GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Harlocker, Susan L.


```

; SEQ ID NO 156
; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-108-605-156

Query Match      12.4%   Score 316.8;   DB 13;   Length 2657;
Best Local Similarity 60.9%;   Pred. No. 6.6e-73;
Matches 570;   Conservative 1;   Mismatches 353;   Indels 12;   Gaps 3;

QY      588  TTCCTACTCAAGACCCAGGACCAACCCCACTTCCTCCCAATTACTAGGGTTGAAGCCATTGC 647
DB      949  TACCACGCACGAGGCTGAGATAACAACTCATGCCATTGCTCGAACACCGTCCCATTC 1008
QY      648  AGCTGTTAGAGTGAACACGAGGGAAGATTGGTTGTCTGGAAGACCCAGTTGCTCA 707
DB      1009  AGCTGCTGGAACAGAGGCCAGTGGTAGATTCCGTTGATGTGTGGCAGCCAGCTCAACA 1068
QY      708  ATGAATATGCGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTCTGGCAGAAAT 767
DB      1069  ATCAGGATGTGGCCGTCAAGATCTTTTCGATCAGGAGAAAAGAAATCGTGGACCCAGC 1128
QY      768  ATGAAGTCTATAGTCTACCTCGAAATGAAGCATGAGAACATATCTACAGTTCAATGGT 827
DB      1129  AGGATATCTACAGCTGCGCGCATCGGCCATCCGAACATCTCTCGAATTCCTGGCGCT 1188
QY      828  AGAAAGAGGACCAAGTGTGGATGTGGACCTGTGGCTAATCACAGCATTTTCATGAAA 887
DB      1189  AGAA-----GCACATGGAACAAGCCGGAATATTGGCTGATATCCACTACAGCATAAC 1242
QY      888  GGTCACTGTCAACATTTCTTAAGGCTAATGTGTGCTCTCTTGGAAATCACTTTGTCA 947
DB      1243  GATCACTATGCGACTACCTCAAAATCGCACAGATCTCATGGCCAGAGTTGTGCCGCA 1302
QY      948  CAGAAACCATGGCTAGAGGATTGGCATATTTACATAGGATATACCTGGCTTAA--AG 1004
DB      1303  CTGAGTCCATGGCCATGAGCTGGCACTTCGACGAGGAGATCCCGGCATCAAGACCG 1362
QY      1005  ATGGCCACAAGCCTGCAATCTCTCACGGGACATCAAAAGTAAATGTGCTTTGAAAA 1064
DB      1363  ATGGGCTAAACCATCGATAGCTCACCGAGACTTCAAGTCTAAGACTTACGCTTAAGA 1422
QY      1065  ACAATCTGACAGCTTGCAATTTGACTTTGGTTGGCTTAAAGTTTCAGAGCTGCAAG 1124
DB      1423  GCGATCTGACGGGCTGTATAGCTGACATTGGTTTGGCCATGATATTCAGCCAGGCA 1482
QY      1125  CTGCAGGTGACCCCATGGCGAGGTTGGTACCCGGAGGTATATGGCTCCAGAGGTGTT 1184
DB      1483  CTTGGCGCGATACACACGGTCAAGTAGGCACTCGACGTTTACATGCCCCAGAGGTC 1542
QY      1185  AGGTTGCTATAAACTTCCAAAGGAGCAATTTCTGAGATAGATATGTCGCCATGGGAT 1244
DB      1543  AGGTTGCCATCAATTTCAATAGAGACGCTTTCTTACGCATAGACGCTACCGCATCG 1602
QY      1245  TAGTCTATGGGAATTTGCTTCTGTTGCACTGTGTCAGATGAGACCCGTTAGATGAG 1304
DB      1603  TAGTCTCTGCGAAATGGTGTACCGGTG---TGACTTTCCCGACCCGTCGTTGAGTTC 1659
QY      1305  TGTTACCACTTTGAGGAGAAATTTGGCCAGCATCTCTTGAAGATATGCGAAGATTG 1364
DB      1660  AGCTGCCCTTTTGAGGCCGAGCTGGGGCTGAGGCCCTGCTGGACGAAGTTTCAGGAG 1719
QY      1365  TTGTGCATAAAAAAAGAGGCTGTTTTAAGAGATTTATGGCGAACAATGCAAGGATG 1424
DB      1720  TGGTAAATGAAGATGCGCCCTCGTTTGTCTCAACTCTCGGCGCGCCATCCCGGAT 1779
QY      1425  CAATGCTCTGTGAACGATAGAGAAATGTTGGATCATGATGACAGAGCCAGGTTATC 1484
DB      1780  ATGTATTCTCGACACAATATGAGGAGTGTCTGGATCACGACGCTGAGGCTCTTAC 1839
QY      1485  CTGGATGTGTAGGTGAAGAAATTTACTTCAGATGCAAA 1520
DB      1840  CTTGCTGTGTATGGAACCGTTTGGCAGCTAACA 1875

```

APPLICANT: Secretist, Heather
 APPLICANT: Wang, Aijun
 APPLICANT: Stolk, John A.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
 FILE REFERENCE: 210121-527C2
 CURRENT APPLICATION NUMBER: US/10/146,502
 CURRENT FILING DATE: 2002-05-14
 NUMBER OF SEQ ID NOS: 2241
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 688
 LENGTH: 452
 TYPE: DNA
 ORGANISM: Homo sapiens
 JS-10-146-502-688

Query Match 15.7%; Score 403.6; DB 14; Length 452;
 Best Local Similarity 93.1%; Pred. No. 1.8e-96;
 Matches 421; Conservative 1; Mismatches 30; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| 2Y | 583 | ACTTGTTCCTACTCAAGACCCAGGACCCACCCACCTTCCCCATTACTAGGGTTGAAGCC | 642 |
| DB | 1 | ACTTGTTCCTACTCAAGACCCAGGACCCACCCACCTTCCCCATTACTAGGGTTGAAGCC | 60 |
| 3Y | 643 | ATTGCAGCTGTAGAGTGAAGACGAGGGAAGATTGGTTGCTCTGGAAGCCAGTT | 702 |
| DB | 61 | ACTGCAGTTATTAGAGTGAAGACGAGGGAAGATTGGTTGCTCTGGAAGCCAGTT | 120 |
| 3Y | 703 | GCTCAATGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGCAGAA | 762 |
| DB | 121 | GCTTAAACGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGCAGAA | 190 |
| 3Y | 763 | TGAATATGAAGTCTATAGTCTACTTGAATGAAGCATGAGAACATACACTACAGTTCAATGG | 822 |
| DB | 181 | TGAATACGAAGTCTACAGTTTGGCTTGAATGAAGCATGAGAACATATTTACAGTTCAATGG | 240 |
| 3Y | 823 | TGCAGAGAAAACAGGACCCAGTGTGGATGTGGCTGTGGCTTAATCAGACATTTTCATGA | 882 |
| DB | 241 | TGCAGAAAACAGGACCCAGTGTGGATGTGGCTTTTGGCTGATCAGAGCATTTTCATGA | 300 |
| 3Y | 883 | AAAGGGCTCACTGTCTAGACATTTCTTAAGGCTTAATGTGGTCTCTTGGAAATCARTTTGTCA | 942 |
| DB | 301 | AAAGGGTTCACATATCAGACATTTCTTAAGGCTTAATGTGGTCTCTTGGAAATCARTTTGTCA | 360 |
| 3Y | 943 | TATTGCAGAAACCATGGCTAGAGGATTGGCATATTTACATGAGGATATACCTGGCTTAAA | 1002 |
| DB | 361 | TATTGCAGAAACCATGGCTAGAGGATTGGCATATTTACATGAGGATATACCTGGCTTAAA | 420 |
| 3Y | 1003 | AGATGCCCAACAGCCTGCAATCTCTCAGGG | 1034 |
| DB | 421 | AGATGCCCAACAGCCTGCAATCTCTCAGGG | 452 |

RESULT 8
 US-10-108-605-156
 ; Sequence 156, Application US/10108605
 ; Publication No. US20020160934A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Broadus, Julie
 ; APPLICANT: Stam, Lynn
 ; APPLICANT: Bachmann, Jane
 ; APPLICANT: Kamdar, Kjm
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
 ; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
 ; FILE REFERENCE: 31133B
 ; CURRENT APPLICATION NUMBER: US/10/108,605
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: US 09/761,142
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/176,418
 ; PRIOR FILING DATE: 2000-01-14
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: PatentIn Ver. 2.1

RESULT 9

10-108-605-78
 Sequence 78, Application US/10108605
 Publication No. US20020160934A1
 GENERAL INFORMATION:
 APPLICANT: Broadus, Julie
 APPLICANT: Stam, Lynn
 APPLICANT: Bachmann, Jane
 APPLICANT: Kamdar, Kim
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
 FILE REFERENCE: 31133B
 CURRENT APPLICATION NUMBER: US/10/108,605
 CURRENT FILING DATE: 2002-03-27
 PRIOR APPLICATION NUMBER: US 09/761,142
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/176,418
 PRIOR FILING DATE: 2000-01-14
 NUMBER OF SEQ ID NOS: 361
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 78
 LENGTH: 2687
 TYPE: DNA
 ORGANISM: Drosophila melanogaster
 S-10-108-605-78

Query Match 12.4%; Score 316.8; DB 13; Length 2687;
 Best Local Similarity 60.9%; Pred. No. 6.7e-73;
 Matches 570; Conservative 1; Mismatches 353; Indels 12; Gaps 3;

Y 588 TTCTACTCAAGACCCAGGACCCACCCACCTTCCCACTTACTAGGTTGAAGCCATTGC 647
 b 979 TACCCAGCAGGCTGAGATAACAACACTCATGCCATTGCTCAACACCGTCCCAATC 1038
 Y 648 AGCTGTAGAGTGAAGCAAGGAGATTTGGTGTCTGGAAGCCCAAGTTGCTCA 707
 b 1039 AGCTGTGGAACAGAACGCCAGTGTAGATTGGTGTGATGTGGCAAGCCCAAGCTCAACA 1098
 Y 708 ATGATATGTGGTGTCAAAATATTTCCAAATACAGGACAAACAGTCTTGGCAGATGAAT 767
 b 1099 ATCAGGATGTGGCGTCAAGATCTTTCCATCGACGAAAAAGAAATCGTGACCAAGC 1158
 Y 768 ATGAAGTCTATAGTCTACTCGAATGAAGCATGAGAACATATACAGTTCATTTGTCATATTG 827
 b 1159 ACGTATCTCAAGTGTGGCGATGGCCATCCGAACATCTCGAATCTCTGGCGGTTG 1218
 Y 828 AGAAAAGAGGACCAAGTGTGGATGTGGACCTGTGGCTATATCAGCATTTTCATGAAAGG 887
 b 1219 AGAA-----GCACATGGACAAAGCGGAATATTGGCTGATATCCACCTACCAAGATAACG 1272
 Y 888 GCTCACTGTAGACTTTCTTAAGGCTAATGTGGTCTCTTGGATCACTTTGTTCATATTG 947
 b 1273 GATCACTATGGACTACTCTCAATCGCACACGATCTCATGGCCAGAGTTGTGGCGCATCG 1332
 Y 948 CAGAAACATGCTAGAGATTGGCATATTTACATGAGGATATACCTGGCTTAAA---AG 1004
 b 1333 CTGAGTCCATGGCCATGAGTGGACATCTGACGAGGAGATCCCGGCATCAAGACCG 1392
 Y 1005 ATGGCCCAAGGCTGCAATCTCTCAGAGGACATCAAAAGTAAATATGCTGTGAAAA 1064
 b 1393 ATGGGCTAAACCATCATGATAGTCAACGAGACTTCAAGTCTAAGAACGTACTGCTTAAGA 1452
 Y 1065 ACAATCTGACAGTTCGATTTGCTGACTTTGGTGGCTTAAAGTTGAGCTGGCAGCT 1124
 b 1453 GCGATCTGACGGCTCTGATAGCTGACTTTGGTTGGCCATGATATTCAGCCAGGCAAGC 1512
 Y 1125 CTGAGGTGACACCCATGGCAGGTTGGTACCGAGGTATATGCTCCAGAGGTGTTGG 1184
 b 1513 CTGCGCGCATACACAGGTCAGTAGGCACTGACCGTTATGCGCCCAAGGTTGCTTG 1572
 Y 1185 AGGGTGTATAAATCTTCCAAAGGACGCAATTTCTGAGGATAGATATGTACGCCATGGAT 1244

RESULT 10

US-09-918-995-30472
 ; Sequence 30472, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 30472
 ; LENGTH: 535
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(535)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-918-995-30472

Query Match 9.3%; Score 238.2; DB 10; Length 535;
 Best Local Similarity 71.1%; Pred. No. 1.7e-52;
 Matches 328; Conservative 1; Mismatches 129; Indels 3; Gaps 1;

Y 784 ACCTGGAATGAAGCATGAGACATACACTACAGTTTCATTGGTGCAGAGAAAAGAGCCAG 843
 b 75 ACCTGGCATGAAGCACGAGAACCTCTACAGTTTCATTGGTGCAGAGAGAGGCTCAA 134
 Y 844 TGTGATGTGGACCTGTGGCTAATCAGCATTTTCATGAAAGGGCTCACTGTCAAGCTT 903
 b 135 COTCAAGTAGAGCTGTGGCTCATTCAGGCTTCCATGACAGGGCTCCCTCAGGATTA 194
 Y 904 TCTTAAGGCTAATGTGGTCTCTTGGAAATCACTTTGTCAATTGTCAGAAACCATGGTAG 963
 b 195 CTTCAAGGGGAACATCATCACATGGAACGACTGTGTGTCATGTAGCAAAAGACGATGTCA 254
 Y 964 AGGATTGGCATATTTACATGAGGATATACC---TGGCTTAAAGATGCGCAAGCCCTGC 1020
 b 255 AGGCTCTCATACCTGATGAGGATGTGCCCTGTGGTGGGAGGCGCAAGCCCTGC 314
 Y 1021 AATCTCTCAGGACATCAAAAGTAAATATGCTGTGAAAAACAATCTGACAGCTTG 1080
 b 315 TATTGCCCCAGGAGCTTTAAAAAGTAAAGTATTTGCTGAAAGGACCTCACAGCGCT 374

TITLE OF INVENTION: MUTATION OF RECEPTOR
FILE REFERENCE: 062361.0108
CURRENT APPLICATION NUMBER: US/09/878,905
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: 08/417,867
PRIOR FILING DATE: 1995-04-07
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 2090
TYPE: DNA
ORGANISM: human

NAME/KEY: CDS
LOCATION: (336)..(2036)

S-09-878-905-10

Query Match 6.6%; Score 168.2; DB 9; Length 2090;
Best Local Similarity 53.8%; Pred. No. 1.7e-33;
Matches 441; Conservative 0; Mismatches 363; Indels 15; Gaps 4;

Y 709 TGAATATGCTGCTCAAAATATTTCCAAATACAGGACAAACAGTCCCTGGCAGAAATGAATA 768
b 1148 TGAGACAGTGGCAGTCAAGATCTTTCCCTATGAGGATATGCCTCTTGGAAAGACAGAGAA 1207

Y 769 TGAAGTCTATAGTCTACCTGGAATGAAGCATGAGACATACATCAAGTTTCATTTGGTGCAGA 828
b 1208 GGACATCTTCTCAGACATCAATCTGAGCATGAGAACATACCTCCAGTTCTCTGACGGCTGA 1267

Y 829 GAAAGAGGACCAAGTGTGATGTGAGCTGTGGGTATATCAAGCATTTTCATGAAAGGG 888
b 1268 GGAGCGGAGACGAGTGTGGGAAACAATACTGGCTGATCACCGCTTCCAGCGCAAGGG 1327

Y 889 CTCAGTGTGAGCTTTCTTAAGGCTATAGTGTCTTGGGAATCACTTTTGTTCATATATGC 948
b 1328 CAACCTACAGGAGTACCTGACCGCGCATGTATCAGCTGGGAGACCTCGGCAAGCTGG 1387

Y 949 AGAAACCATGGCTAGAGATGGCATATTTACATGAGGATATACCTGGCTTAAAGATGG 1008
b 1388 CAGCTCCCTCGCCGGGGATGTCTACCTCCACAGTGTACACT---CCATGTGGGAG 1444

Y 1009 CCACAGCTGCAATCTCTCACAGGACATCAAAAGTAAATATGTCTTTGAAAGACAA 1068
b 1445 GCCCAAGTGCCTATCGTGCACAGGACCTCAAGAGCTCCAATATCTCTGTAAGAACGA 1504

Y 1069 TCTGACAGCTTGCATGTGCTGCTTTGGTGGCTTAAAGTTCAGGCTGGCAAGTCTGC 1128
b 1505 CCTAACCTGCTGCTGTGACTTTGGGCTTTCCCTCGCTCTGGACCTACTCTGTCTGT 1564

Y 1129 AGGTGAC-----ACCCATGGCAGGTTGGTACCGGAGTATATGCTCCAGAGTGT 1182
b 1565 GGATGACCTGGCTAACAGTGGCAGGTGGAACTGCAAGATACATGGCTCCAGAGTCT 1624

Y 1183 GGAGGGTGTATAAACTTCCAAA---GGACGCTATTTCTGAGGATAGATATGACGCCAT 1239
b 1625 AGAATCCAGGATGAATTTGGAGAAATGCTGAGTCTTCAAGCAGACCGATGTCTACTCCAT 1684

Y 1240 GGGATTAGTCTATGGAAATTTGGCTTCTGTTGCACTGCTGCAAGTGCAGACCGTAGATGA 1299
b 1685 GGCTCTGGTGTCTGGAAATGACATCTGCTG---TAATGCACTGGGAGAGTAAAGA 1741

Y 1300 GTACATGTTACATTTGAGGAGAAATTTGGCAGCATCTCTTGAAGATATGACGGA 1359
b 1742 TTATGAGCTCCATTTGGTTCCAAAGTGGGGAGACCCCTGTGTGCGAAAGCATGAAGGA 1801

Y 1360 AGTTGTTGCTATAAAAGAGGCTGTTTAAGAGATTTATGGCAGAAACATGACAGG 1419
b 1802 CAACGTGTTGAGAGATCGAGGGGACACAGAAATTTCCAGCTTCTGGCTCAACCCAGGG 1861

Y 1420 AATGGCAATGCTCTGTGAAACGATAGAGAAATTTGGGATCATGATGACAGAGCCAGGTT 1479
b 1862 CATCCAGATGGTGTGAGACGTTGACTGAGTGTGGGACCCAGACCCAGAGGGCCGCTCT 1921

RESULT 14

US-10-101-510-18
; Sequence 18, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 2090
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-18

Query Match 6.6%; Score 168.2; DB 14; Length 2090;
Best Local Similarity 53.8%; Pred. No. 1.7e-33;
Matches 441; Conservative 0; Mismatches 363; Indels 15; Gaps 4;

Y 709 TGAATATGCTGCTCAAAATATTTCCAAATACAGGACAAACAGTCCCTGGCAGAAATGAATA 768
b 1148 TGAGACAGTGGCAGTCAAGATCTTTCCCTATGAGGATATGCCTCTTGGAAAGACAGAGAA 1207

Y 769 TGAAGTCTATAGTCTACCTGGAATGAAGCATGAGACATACATCAAGTTTCATTTGGTGCAGA 828
b 1208 GGACATCTTCTCAGACATCAATCTGAGCATGAGAACATACCTCCAGTTCTCTGACGGCTGA 1267

Y 829 GAAAGAGGACCAAGTGTGATGTGAGCTGTGGGTATATCAAGCATTTTCATGAAAGGG 888
b 1268 GGAGCGGAGACGAGTGTGGGAAACAATACTGGCTGATCACCGCTTCCAGCGCAAGGG 1327

Y 889 CTCAGTGTGAGCTTTCTTAAGGCTATATGTGTCTTCTTGGAAATCACTTTTGTTCATATATGC 948
b 1328 CAACCTACAGGATACCTGACCGGCATGTATCAGCTGGGAGGACCTCGGCAAGCTGG 1387

Y 949 AGAAACCATGGCTAGAGATTTGGCATATTTACATGAGGATATACCTGGCTTAAAGATGG 1008
b 1388 CAGCTCCCTCGCCGGGGATGTCTACCTCCACAGTGTACACT---CCATGTGGGAG 1444

Y 1009 CCACAGCTGCAATCTCTCACAGGACATCAAAAGTAAATATGTCTTTGAAAGACAA 1068
b 1445 GCCCAAGATGCCATCGTGCACAGGACCTCAAGAGCTCCAATATCTCTGTAAGAACGA 1504

Y 1069 TCTGACAGCTTGCATGTGCTGCTTTGGTGGCTTAAAGTTCGAGGCTGGCAAGTCTGC 1128
b 1505 CCTAACCTGCTGCTGTGACTTTGGGCTTTCCCTCGCTCTGGACCTACTCTGTCTGT 1564

Y 1129 AGGTGAC-----ACCCATGGCAGGTTGGTACCGGAGTATATGCTCCAGAGTGT 1182
b 1565 GGATGACCTGGCTAACAGTGGCAGGTGGAACTGCAAGATACATGGCTCCAGAGTCT 1624

Y 1183 GGAGGGTGTATAAACTTCCAAA---GGACGCTATTTCTGAGGATAGATATGACGCCAT 1239
b 1625 AGAATCCAGGATGAATTTGGAGAAATGCTGAGTCTTCAAGCAGACCGATGTCTACTCCAT 1684

Y 1240 GGGATTAGTCTATGGAAATTTGGCTTCTGTTGCACTGCTGCAAGTGCAGACCGTAGATGA 1299
b 1685 GGCTCTGGTGTCTGGAAATGACATCTGCTG---TAATGCACTGGGAGAGTAAAGA 1741

Y 1300 GTACATGTTACATTTGAGGAGAAATTTGGCAGCATCTCTTGAAGATATGACGGA 1359
b 1742 TTATGAGCTCCATTTGGTTCCAAAGTGGGGAGACCCCTGTGTGCGAAAGCATGAAGGA 1801

1360 AGTTGTTGTCATAAAAAAAGAGGCTGTTTAAAGAGATTATTGGCAGAAACATGACGG 1419
1802 CAACGTTTGAAGATCGAGGGCGACAGAAATTCAGCTTCTGGCTCAACACCAAGGG 1861
1420 AATGGCAATCTCTGTGAAACGATAGAAAGATGTTGGATCATGATGCAAGACAGGTT 1479
1862 CATCCAGATGGTGTGTGAGACGTTGACTGAGTGTGGGACCAACGACCCAGAGGCCGCT 1921
1480 ATCAGCTGATGTAGTGAAGAAATTAATCAATGCA 1518
1922 CACAGCCCACTGTGTGGCAGAAACGCTTCAGTGAGCTGGA 1960

RESULT 15
US-10-101-510-438
Sequence 438, Application US/10101510
Publication No. US20030148295A1
GENERAL INFORMATION:
APPLICANT: WAN, JACKSON
APPLICANT: WANG, YIXIN
TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
FILE REFERENCE: 15117.0012
CURRENT APPLICATION NUMBER: US/10/101.510
CURRENT FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: 60/276,947
PRIOR FILING DATE: 2001-03-20
NUMBER OF SEQ ID NOS: 805
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 438
LENGTH: 3206
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (2351)
OTHER INFORMATION: a, t, c, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (2362)
OTHER INFORMATION: a, t, c, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (2368)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-438

Query Match 6.8%; Score 168.2; DB 14; Length 3206;
Best Local Similarity 53.8%; Pred. No. 2:2e-33;
Matches 441; Conservative 0; Mismatches 363; Indels 15; Gaps 4;
709 TGAATATGTGGCTGTCAAATATTTCCAAATACAGGACAAACAGTCCTGGCAGAAATGAATA 768
1151 TGAGACAGTGGCAGTCAAGATCTTTCCCTATGAGGAGTATGCCCTTTTGGAAAGACAGAGAA 1210
769 TGAAGTCTATAGTCTACCTCGAATGAAGATGAAGAAATATCACTACAGTTCTTGGTGCAGA 828
1211 GGACATCTTCTCAGACATCAATCTGAAGCATGAAGAAATATCACTACAGTTCTTGGTGCAGA 1270
829 GAAAGAGGACCAAGTGTGAGATGGACCTGTGCTATCACTACAGATTTTCATGAAGAGG 888
1271 GGAGCGGAGACGAGTGTGGGAAACAAATCTGGCTGATCAACCGCTTCCAGCCCAAGGG 1330
889 CTCACTGTCAAGCTTTCTTAAGGCTATGTGCTCTCTTGGAAATCACTTTTGTCAATATTGC 948
1331 CAACCTACAGGAGTACCTGACGCGGATGTCAATAGTGGGAGACCTGGCGAAGCTGGG 1390
949 AGAACCATGCTAGAGATTGGGATATTTACATGAGGATATACCTGGCTTAAAGATGG 1008
1391 CAGTCTCTCCCGGGGGATGTCTCACTCCACAGTATCACT---CCATGTGGGAG 1447
1009 CCACAAGCCTGCAATCTCTCAGGAGCAATCAAAAGTAAATATGCTGTTGAAACAA 1068

1448 GCCCAAGATGCCATCGTGTGACAGGACCTCAAGAGCTCAATATCTCTGTGAAGACGA 1507
1069 TGTGACAGCTTGCATTTGCTGACTTTTGGGTTGGCTTTAAAGTTGAGGCTGGCAAGTCTGC 1128
1508 CCTAACCTCTGCTGCTGTGACCTTTGGGCTTTCCCTGCGTCTGAGACCTACTCTGTCTGT 1567
1129 AGGTGAC-----ACCCATGGGAGGTTGGTACCCGAGGATATATGGTCCAGAGGCTGT 1182
1568 GATGACCTGGCTAACAGTGGGAGGTGGAACTGCAAGATACATGGTCCAGAGTCTCT 1627
1183 GGAGGCTGTATAAATTCCTCAAA---GGGACGCAATTTCTGAGGATAGATATGTACGCCAT 1239
1628 AGAATCCAGGATGAATTTGGAGAATGTTGAGTCTTCAAGCAGACCCGATGTCTACTCCAT 1687
1240 GGGATTAGTCTATGGGAATTTGCTTCTCGTTGCACTGCTGAGATGGACCCGATAGATGA 1299
1688 GGTCTGTGGTCTCTGGGAAATGACATCTCGTG---TAAATGACGTGGGAGAGTAAAGCA 1744
1300 GTACATGTTTACCATTTGAGGAGAAATTTGGTCCAGGAGTGGGAGACCCCTGTGTGCAAGCATGAAGCA 1804
1745 TTATGAGCCTCCTCAATTTGGTCCAGGAGTGGGAGACCCCTGTGTGCAAGCATGAAGCA 1864
1360 AGTTGTTGTGCATAAAAAAAGAGGCTGTTTTAAAGATTTATTGGCAGAAACATGCGAG 1419
1805 CAACGTTTTCAGAGATCGAGGCGACCAAGAAATTTCCAGCTTCTGGCTCAACCCAGGG 1864
1420 AATGGCAATGCTCTGTGAAACGATAGAGATGTTGGGATCATGATGAGAAAGCCAGGTT 1479
1865 CATCCAGATGGTGTGTGAGACGTTGACTGAGTGTGGGACCAAGACCCAGAGGCCGCTCT 1924
1480 ATCAGCTGGATGTGTAGTGAAGAAATTAATCACTCAGATGCA 1518
1925 CACAGCCCACTGTGTGGCAGAAACGCTTCAGTGAGCTGGA 1963

Search completed: February 26, 2004, 08:15:33
Job time : 3111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

4 nucleic - nucleic search, using sw model

in on: February 26, 2004, 02:39:37 ; Search time 5885 Seconds
(without alignments)
13005.398 Million cell updates/sec

itle: US-09-742-684A-15
erfect score: 2563
equence: 1 ctctcggagagaccagggga.....aacacgctttacaataagcc 2563

coring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

earched: 27513289 seqs, 14931090276 residues

otal number of hits satisfying chosen parameters: 55026578

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_esthum:*
4: em_esthum:*
5: em_esthum:*
6: em_esthum:*
7: em_esthum:*
8: em_esthum:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estfun:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdt:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2430.6 | 94.8 | 4533 | 11 | AK041246 Mus muscu |
| 2 | 1303.4 | 50.9 | 1427 | 29 | AY407987 Mus muscu |
| 3 | 1300.4 | 50.7 | 1542 | 29 | AY407985 Homo sapi |
| 4 | 1020.6 | 39.8 | 1487 | 29 | AY407986 Pan trogl |

| | | | | | |
|----|-------|------|------|----|----------|
| 5 | 782.8 | 30.5 | 801 | 14 | CF743932 |
| 6 | 737.6 | 28.8 | 779 | 14 | CD642673 |
| 7 | 726.6 | 28.3 | 899 | 13 | BQ960311 |
| 8 | 721.2 | 28.1 | 849 | 12 | BI827084 |
| 9 | 705 | 27.5 | 717 | 14 | CB245658 |
| 10 | 701.4 | 27.4 | 728 | 13 | BU704233 |
| 11 | 695.8 | 27.1 | 742 | 14 | CK031062 |
| 12 | 668.8 | 26.1 | 763 | 14 | CF271903 |
| 13 | 668.8 | 26.1 | 776 | 14 | CK032821 |
| 14 | 662.4 | 25.8 | 786 | 14 | CF271908 |
| 15 | 661.2 | 25.8 | 1487 | 29 | AY421275 |
| 16 | 658.2 | 25.7 | 816 | 14 | CK032823 |
| 17 | 654.6 | 25.5 | 777 | 14 | CK130200 |
| 18 | 647.4 | 25.3 | 797 | 14 | CF271905 |
| 19 | 645.8 | 25.2 | 802 | 14 | CF271907 |
| 20 | 640.8 | 25.0 | 807 | 14 | CK032824 |
| 21 | 639 | 24.9 | 796 | 14 | CK032822 |
| 22 | 634.8 | 24.8 | 763 | 14 | CF271906 |
| 23 | 623.8 | 24.3 | 1487 | 29 | AY421277 |
| 24 | 623.4 | 24.3 | 626 | 29 | CG425408 |
| 25 | 619.6 | 24.2 | 765 | 14 | CK130198 |
| 26 | 614.8 | 24.0 | 700 | 13 | BX485334 |
| 27 | 608.8 | 23.8 | 742 | 14 | CF271904 |
| 28 | 602.6 | 23.5 | 1045 | 12 | BI558693 |
| 29 | 597.4 | 23.3 | 610 | 28 | AZ900935 |
| 30 | 590.8 | 23.1 | 774 | 12 | BG419580 |
| 31 | 546.8 | 21.3 | 667 | 14 | CF108155 |
| 32 | 537 | 21.0 | 648 | 13 | BQ211132 |
| 33 | 533.4 | 20.8 | 602 | 14 | CB142878 |
| 34 | 529 | 20.6 | 945 | 14 | CA985889 |
| 35 | 518.2 | 20.2 | 930 | 14 | CF407213 |
| 36 | 514.2 | 20.1 | 741 | 13 | BU451326 |
| 37 | 506 | 19.7 | 801 | 13 | BU747194 |
| 38 | 503.6 | 19.6 | 582 | 10 | BF044914 |
| 39 | 496.6 | 19.4 | 747 | 13 | BU370235 |
| 40 | 489.8 | 19.1 | 937 | 14 | CF407212 |
| 41 | 485.6 | 18.9 | 763 | 14 | CF408089 |
| 42 | 477.4 | 18.6 | 709 | 14 | CF271902 |
| 43 | 472.2 | 18.4 | 759 | 9 | AI829731 |
| 44 | 470.8 | 18.4 | 960 | 14 | CA982309 |
| 45 | 463.2 | 18.1 | 1365 | 29 | AY421276 |

ALIGNMENTS

RESULT 1
AK041246
LOCUS
DEFINITION
Mus musculus adult male aorta and vein cDNA, RIKEN full-length enriched library, clone: A530094G04 product: activin receptor IIA, full insert sequence.
4533 bp mRNA linear HTC 19-SEP-2003
ACCESSION
AK041246
VERSION
AK041246.1 GI:26334318
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159

REFERENCE
AUTHORS

3 Shibata,K., Itoh,M., Aizawa,K., Nagao,K., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

TITLE
JOURNAL
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE
JOURNAL
REFERENCE
AUTHORS

6 (bases 1 to 4533)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

TITLE
JOURNAL

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 4533
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:A530094G04"
/db_xref="MGI:2406272"
/db_xref="taxon:10090"
/clone="A530094G04"
/sex="male"
/tissue_type="aorta and vein"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
617. 2158
/note="unnamed protein product; activin receptor IIA
(MGD|MG1:102806, GB|NM_007396, evidence: BLASTN, 99%,
match=2477)
putative"
/codon_start=1
/protein_id="BAC30877.1"

COMMENT

FEATURES
source

1. 4533
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:A530094G04"
/db_xref="MGI:2406272"
/db_xref="taxon:10090"
/clone="A530094G04"
/sex="male"
/tissue_type="aorta and vein"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
617. 2158
/note="unnamed protein product; activin receptor IIA
(MGD|MG1:102806, GB|NM_007396, evidence: BLASTN, 99%,
match=2477)
putative"
/codon_start=1
/protein_id="BAC30877.1"

CDS

1. 4533
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:A530094G04"
/db_xref="MGI:2406272"
/db_xref="taxon:10090"
/clone="A530094G04"
/sex="male"
/tissue_type="aorta and vein"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
617. 2158
/note="unnamed protein product; activin receptor IIA
(MGD|MG1:102806, GB|NM_007396, evidence: BLASTN, 99%,
match=2477)
putative"
/codon_start=1
/protein_id="BAC30877.1"

/db_xref="GI:26334319"
/translation="MGAAKLAFAVFLISSGAILRSETQECLEFNANWERDRTNQ
TGPEYCGDKRKHCFATWKNISGFIEIVKQGWLDINDICDRDCEKSDSEVYF
CCCEGNMCKEYSYPPEMEVTPQSPNPVTPKPYINILLVSLVPLMILAGIVICAFW
YHHRKAYPPVLPVTPDGPSPPLGLKPLQLLEVKARGFQGWCAQLINLEYAV
KLPIPOQKSWNEIYEVYSLPGMKHENILOFIEGAEKGTSDVDLWLLTAPEHKSLS
DLKANVSWNEILHAETWARGLYLHEIPGLKDGHPALSDHEDIKSNVLLKNL
TACIADFGALKFEAGKSGADTHGVGTRRMAPEVLEAGINLORDAFLRIDMYAMGL
VLMELASRCTAADGPVDEYMLPEEEIGQHPHSDLEDMQEVVHKKRPVLRDWTQKHG
MANLCTIEBCWDHDAEARLSAGCVGERITQMRITNIITTEDIVTVVMTVNTVDFPP
KESSL"

ORIGIN

| | Query Match | Score | DB 11; | Length | 4533; |
|----|-----------------------|--|---------|------------|--------------|
| | Best Local Similarity | 94.8%; | 2430.6; | DB 11; | Length 4533; |
| | Matches 2497; | Conservative | 2; | Mismatches | 57; |
| | | | | Indels | 4; |
| | | | | Gaps | 3; |
| QY | 1 | CTCCGAGGAGACCCAGGAACTGATATCTAGCGAGAACTCTTACGGCTTCTCCGGCG | 60 | | |
| DB | 547 | CTCCGAGGAGACCCAGGAACTGATATCTAGCGAGAACTCTTACGGCTTCTCCGGCG | 606 | | |
| QY | 61 | CTCCGAGGAAATGGAGCTGCTGCAAGTGGCGTTGGCGTCTTCTTATCTTCTGCTC | 120 | | |
| DB | 607 | CTCCGAGGAAATGGAGCTGCTGCAAGTGGCGTTGGCGTCTTCTTATCTTCTGCTC | 666 | | |
| QY | 121 | TTCAGGTCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTAATGCTAATTG | 180 | | |
| DB | 667 | TTCAGGTCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTAATGCTAATTG | 726 | | |
| QY | 181 | GGAAAGACAGAGAACCAACAGCTGGTGTGAACCTTGTATGGTATGAAGATAACG | 240 | | |
| DB | 727 | GGAAAGACAGAGAACCAACAGCTGGTGTGAACCTTGTATGGTATGAAGATAACG | 786 | | |
| QY | 241 | GGCAGATTTTGTCTGCTGAGAGATTTCTGCTTCCATTTCCATTTGAAGCAAGG | 300 | | |
| DB | 787 | GGCAGATTTTGTCTGCTGAGAGATTTCTGCTTCCATTTCCATTTGAAGCAAGG | 846 | | |
| QY | 301 | TTGTGGCTGATGATATCACTGATGACAGCACTGATTTGTGTGAGAAAGAACAG | 360 | | |
| DB | 847 | TTGTGGCTGATGATATCACTGATGACAGCACTGATTTGTGTGAGAAAGAACAG | 906 | | |
| QY | 361 | CCCTGAAGTGTACTTTTGTGTGTGAGGCAATATGTGTAAGAAAGTTCTTATT | 420 | | |
| DB | 907 | CCCTGAAGTGTACTTTTGTGTGTGAGGCAATATGTGTAAGAAAGTTCTTATT | 966 | | |
| QY | 421 | TCGGAGATGGAAGTCAACAGCCACTTCAATCTCTTACACCGAGCCACCTATTA | 480 | | |
| DB | 967 | TCGGAGATGGAAGTCAACAGCCACTTCAATCTCTTACACCGAGCCACCTATTA | 1026 | | |
| QY | 481 | CAACATTCCTGTATCTTGTGTGACCACTAATCTTAATTCGAGCAATTTGCTTGC | 540 | | |
| DB | 1027 | CAACATTCCTGTATCTTGTGTGACCACTAATCTTAATTCGAGCAATTTGCTTGC | 1086 | | |
| QY | 541 | ATTTTGGTGTACAGACATCAAGATGGCTACCTCTCTTCTTCTTCTTCTCAAGA | 600 | | |
| DB | 1087 | ATTTTGGTGTACAGACATCAAGATGGCTACCTCTCTTCTTCTTCTTCTCAAGA | 1146 | | |
| QY | 601 | CCAGGACCAACCCACCTTCCCATTTACTAGGTGTGAAGCCATTGCGAGCTTTAGA | 660 | | |
| DB | 1147 | CCAGGACCAACCCACCTTCCCATTTACTAGGTGTGAAGCCATTGCGAGCTTTAGA | 1206 | | |
| QY | 661 | GAAAGCAAGGGGAAGATTGTGTGTGAAAGCCAGTTGCTCAATGAATATGTGCG | 720 | | |
| DB | 1207 | GAAAGCAAGGGGAAGATTGTGTGTGAAAGCCAGTTGCTCAATGAATATGTGCG | 1266 | | |
| QY | 721 | TGTCAAAATATTTCCATACAGGCAAAACAGTCTGGCAAGATGAATGAAGTCTAT | 780 | | |
| DB | 1267 | TGTCAAAATATTTCCATACAGGCAAAACAGTCTGGCAAGATGAATGAAGTCTAT | 1326 | | |
| QY | 781 | TCTACCTGGAATGAGCATGAGAACTACTACAGTTTCTTGTGTCAGAAAGAGGCAC | 840 | | |
| DB | 1327 | TCTACCTGGAATGAGCATGAGAACTACTACAGTTTCTTGTGTCAGAAAGAGGCAC | 1386 | | |

841 CAGTGTGGATGTGGACCTGTGGCTAAATCACAGCATTTTCATGAAAGGGCTCACTGTGAGA 900
1387 CAGTGTGGATGTGGACCTGTGGCTAAATCACAGCATTTTCATGAAAGGGCTCACTGTGAGA 1446
901 CTTTCTTAAGGCTAATGTGGTCTCTTGGAAATCAACCTTTGTCTATTTGCAAGAACCATGGC 960
1447 CTTTCTTAAGGCTAATGTGGTCTCTTGGAAATCAACCTTTGTCTATTTGCAAGAACCATGGC 1506
961 TAGAGGATGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAAGCCTGC 1020
1507 TAGAGGATGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAAGCCTGC 1566
1021 AATCTCTCAGAGGACATCAAAAGTAAATGTGTGTGTGAAACAACTGACAGCTTG 1080
1567 AATCTCTCAGAGGACATCAAAAGTAAATGTGTGTGTGAAACAACTGACAGCTTG 1626
1081 CATTGCTGACCTTGGTGGCTTAAAGTTCGAGGCTGCAAGTCTGCAAGTGGGACACCCCA 1140
1627 CATTGCTGACCTTGGTGGCTTAAAGTTCGAGGCTGCAAGTCTGCAAGTGGGACACCCCA 1686
1141 TGGCAGGCTTGGTACCCGAGGATATATGGCTCCAGAGGTGTTGGAGGGTCTATAAATTT 1200
1687 TGGCAGGCTTGGTACCCGAGGATATATGGCTCCAGAGGTGTTGGAGGGTCTATAAATTT 1746
1201 CCAAGGAGCGCATTTCTGAGGATAGATATGTACGCCATGGGATTTAGTCTATGGGAATT 1260
1747 ACAAGGAGCGCATTTCTGAGGATAGATATGTACGCCATGGGATTTAGTCTATGGGAATT 1806
1261 GCTTCTCGTTCGATGCTGCGAGATGCAAGTGTAGATGATGATGATGATGATGATGATGATG 1320
1807 GCTTCTCGTTCGATGCTGCGAGATGCAAGTGTAGATGATGATGATGATGATGATGATGATG 1866
1321 AGAAATGGCCAGCATCCATCTCTGAGGATATGCAAGGATGTTGTCGATATAAAAAA 1380
1867 AGAAATGGCCAGCATCCATCTCTGAGGATATGCAAGGATGTTGTCGATATAAAAAA 1926
1381 GAGGCTGTTTTAAGAGATTTATGGCAGAAACATGCAAGGATGCAAGTGTCTGTGAAC 1440
1927 GAGGCTGTTTTAAGAGATTTATGGCAGAAACATGCAAGGATGCAAGTGTCTGTGAAC 1986
1441 GATAGAAGAAATGTTGGGATCATGATGAGAGCCAGGTTATCAGTGGATGTAGTGA 1500
1987 GATAGAAGAAATGTTGGGATCATGATGAGAGCCAGGTTATCAGTGGATGTAGTGA 2046
1501 AAGAAATCTCAGATGCAAGATGCAAGATATCATTTACTACAGAGGACATTTGTAACAGT 1560
2047 AAGAAATCTCAGATGCAAGATGCAAGATATCATTTACTACAGAGGACATTTGTAACAGT 2106
1561 GGTCAATGGTGCACAAATGTTGACTTTCCTCCAAAGAACTTCTATGATGTTGCA 1620
2107 GGTCAATGGTGCACAAATGTTGACTTTCCTCCAAAGAACTTCTATGATGTTGCA 2166
1621 CCGTCTGTACACTGAGGACTGGGACTCTGAAGTGGAGTCTGAAGTAAAGAAAGTGC 1680
2167 CCGTCTGTACACTGAGGACTGGGACTCTGAAGTGGAGTCTGAAGTAAAGAAAGTGC 2226
1681 TTAGTGTGATTTCTGTGAAATAGTAGGATGCTCCAGGACATGTACGAGAGGCCCC 1740
2227 TTAGTGTGATTTCTGTGAAATAGTAGGATGCTCCAGGACATGTACGAGAGGCCCC 2286
1741 CTTGTGAAAGCATGGATCTGGGAGATGGATCTGGGAATTTACTGCTCTGAGCA 1800
2287 CTTGTGAAAGCATGGATCTGGGAGATGGATCTGGGAATTTACTGCTCTGAGCA 2346
1801 CAGATATGAAGAGGATCTAAGGGAAAGCTGCAAACTGTAAAGAACTTCTGAAAATGTA 1860
2347 CAGATATGAAGAGGATCTAAGGGAAAGCTGCAAACTGTAAAGAACTTCTGAAAATGTA 2406
1861 CTCGAAGATGTTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAATCAAGTATTT 1920
2407 CTCGAAGATGTTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAATCAAGTATTT 2466
1921 GCAAAACTGACATCAGATTTCTTAAATGTGTGTGAGAGACATTAATTCCTTAATGAACT 1980

2467 G-AAAACCTGACATCAGATTTCTTAAATGTCTGTGAGAGACACTAATTCCTTAAATGAACT 2525
1981 ACTGCTAATTTTTTAAATGAAAACTTTTTCATTTTTCATTTTAAAGGGTAACTTTTT 2040
2526 ACTGCTAATTTTTTAAATGAAAACTTTTTCATTTTTCATTTTAAAGGGTAACTTTTT 2585
2041 ATTGCATTTGCTGTGTTTCTTATTAATGACATATTGTATGCAACATGACAGCTTTGTG 2100
2586 ATTGCATTTGCTGTGTTTCTTATTAATGACATATTGTATGCAACATGACAGCTTTGTG 2645
2101 AATGTGTAGTGTGCTGTTTCTGTGTACATAGTCAAGTGGGGTACAGTAAAGAGG 2160
2646 AATGTGTAGTGTGCTGTTTCTGTGTACATAGTCAAGTGGGGTACAGTAAAGAGG 2705
2161 CTTCCAGCATTTTAACTTAACTCCCTCAACAGGTATACCTCAGTCCACGGTGTGCTAAA 2220
2706 CTTCCAGCATTTTAACTTAACTCCCTCAACAGGTATACCTCAGTCCACGGTGTGCTAAA 2765
2221 TTATAAAATTTGAAAACTAATACAGAAATTTGAATAAATCATGTCATGTTTATTAACAAGT 2280
2766 TTATAAAATTTGAAAACTAATACAGAAATTTGAATAAATCATGTCATGTTTATTAACAAGT 2825
2281 TAATTACAAATTTCACTGTTTATTTAAGAAAAATGTTAAGTATGCTTGTAGTGCCTAATAG 2340
2826 --ATTACAAATTTCACTGTTTATTTAAGAAAAATGTTAAGTATGCTTGTAGTGCCTAATAG 2883
2341 TAAGTGGCTATTTTAAAGCAGTGTGTTTCTTCTTACTGGCTTGTAAATTTAGGGA 2400
2884 TAAGTGGCTATTTTAAAGCAGTGTGTTTCTTCTTACTGGCTTGTAAATTTAGGGA 2943
2401 AAAC--AAGTGTGCTTCTTCAAAATGGAAGAAATATGTTGTCACCTACCCCTACCTA 2459
2944 AAACAAAGTGTGCTTCTTCAAAATGGAAGAAATATGTTGTCACCTACCCCTACCTA 3003
2460 TATCAAGTGTGCTTCTTCAAAATGGAAGAAATATGTTGTCACCTACCCCTACCTA 2519
3004 TATCAAGTGTGCTTCTTCAAAATGGAAGAAATATGTTGTCACCTACCCCTACCTA 3063
2520 ACAAGCAGTGTGCTTCTTCAAAATGGAAGAAATATGTTGTCACCTACCCCTACCTA 2559
3064 TTTGGAAATATTTTGAAGAAATATGGAAGCAATGATTTTAAATTT 3103

RESULT 2
AY407987 1427 bp DNA linear GSS 15-DEC-2003
LOCUS ACVR2 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION Genomic survey sequence.
ACCESSION AY407987
VERSION AY407987.1 GI:39763958
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1427)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL 14671302
PUBMED 2 (bases 1 to 1427)
REFERENCE Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
AUTHORS Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
JOURNAL Rockville, MD 20850, USA

Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..1487
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1487
/genes="ACVR2"
/locus_tag="HCM3073"

Query Match 39.8%; Score 1020.6; DB 29; Length 1487;
Best Local Similarity 71.1%; Pred. No. 3.4e-249;
Matches 1057; Conservative 2; Mismatches 428; Indels 0; Gaps 0;

126 GTGCTATCTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTTGGAAA 185
1 GTGCTATCTGGTAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTTGGAAA 60
186 ARGACAGAACCAACGACTGTGTGTAACCTTGCTATGGTGATAGATAAAGACGCGAC 245
61 AAGACAGAACCAACAACTGTGTGTAACCTTGCTATGGTGATAGATAAAGACGCGAC 120
246 ATTCTTTTGTCTACCTGGAAGAATATTCTGTGTCTTCCATTGAAATAGTGAAGAGTGTGT 305
121 ATTGNN 180
306 GGCTGGATGATATCACTGCTATGACAGAGTGTGTGTGNGAAGAAAAGACAGCCCTG 365
181 NNN 240
366 AAGTGTACTTTTGTGTGTGAGGCAATATGTGTAAATGAAAGTCTCTTATTTCCGG 425
241 NNN 300
426 AGATGGAAGTACACAGCCCACTTCAAACTCTGTACACGGAAGCCACCTTATTAACA 485
301 NNN 360
486 TTCTGTCTATTTCTTGT 545
361 TCTGTCTATTTCTTGT 420
546 GGGGTGTACAGACATCAAGATGGCTTACCTCTCTGTGTGTGTGTGTGTGTGTGTGT 605
421 GGGGTGTACAGGATCAAGATGGCTTACCTCTCTGTGTGTGTGTGTGTGTGTGTGT 480
606 GACCAACCCCACTTCCCACTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
481 GACCAACCCCACTTCCCACTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
666 CAAGGGGAGATTTGGT 725
541 CAAGGGGAGATTTGGT 600
726 AATATTTTCAATACAGGACAAACAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 785
601 AATATTTTCAATACAGGACAAACAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
786 CTGGAATGAAGATGAGAACATCTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 845
661 NNNNNATGAAGATGAGAACATCTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720
846 TGGATGTGACCTGTGGTATCAAGATTTTCAAGAAAGGGCTCACTGTGTGTGTGTGT 905
721 TTGATGTGGATCTTTGGT 780
906 TTAAGGCTAATGTGTCTCTTGGATCARTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 965

Db 781 TTAAGGCTAATGTGTCTCTTGGATGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 966 GATTGGCATTTTACATGAGGATATACCTGGCTTAAAGATGCGCAAGCTTCAATCT 1025
Db 841 GATTGGCATTTTACATGAGGATATACCTGGCTTAAAGATGCGCAAGCTTCAATCT 900
Qy 1026 CTCACAGGAGCATCAAAAGTAAATAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1085
Db 901 CTCACAGNN 960
Qy 1086 CTGACTTTGGTGGCTTTAAAGCTTCGAGCTGGCAAGCTTCGAGGTGACACCCATGGGC 1145
Db 961 NNN 1020
Qy 1146 AGGTTGGTACCCGAGGATATGGCTCCAGAGTGTGGAGGTGCTATAAATTTCCAAA 1205
Db 1021 NNGTTGGTACCCGAGGATATGGCTCCAGAGGTATTAGAGGGTCTATAAATTTCCAAA 1080
Qy 1206 GGGACGATTTCTGAGGATAGATATGAGCCATGGGATAGTCTTATGGGAATTTGGCTT 1265
Db 1081 GGGATGCAATTTTGGAGATAGATATGAGCCATGGGATAGTCTTATGGGAATTTGGCTT 1140
Qy 1266 CTGTTGCACTGTGAGATGGACCCGATAGATGATACATGTTACATTTGAGGAAGAAA 1325
Db 1141 CTGCTGTACTGTGAGATGGACCTGTAGATGAATACATGTTGCAATTTGAGGAGAAA 1200
Qy 1326 TTGGCCAGCATCCATCTCTTGAAGATATGAGGAAGTGTGTGTGTGTGTGTGTGTGTGT 1385
Db 1201 TTGGCCAGCATCCATCTCTTGAAGATATGAGGAAGTGTGTGTGTGTGTGTGTGTGT 1260
Qy 1386 CTGTTTAAAGATATTTGGCAGAAACATGAGGAATGCAATGCTCTGTGAAACGATAG 1445
Db 1261 CTGTTTAAAGATATTTGGCAGAAACATGAGGAATGCAATGCTCTGTGAAACGATAG 1320
Qy 1446 AAGATGTTGGATCATGATGAGAGCCAGGTATCAGCTGATGTGTAGGTGAAAGAA 1505
Db 1321 AAGATGTTGGATCANGACGNNNAGCCAGGTATCAGCTGATGTGTAGGTGAAAGAA 1380
Qy 1506 TTACTCAGATGCAAGACTTACAAATATCATTACTACAGAGCATGTAAACAGTGTGCA 1565
Db 1391 TTACCAAGTGCAGAGACTTACAAATATTTATACACAGAGCATGTAAACAGTGTGCA 1440
Qy 1566 CAATGTTGACAAATGTTGACTTTCTTCCCAAGAAATCTAGTCTATGA 1612
Db 1441 CAATGTTGACAAATGTTGACTTTCTTCCCAAGAAATCTAGTCTATGA 1487

CF743932 801 bp mRNA linear EST 10-OCT-2003
UI-M-G10-clq-d-20-0-UI.r1 NIH BMAP_G10 Mus musculus cDNA clone
IMAGE:30616699 5', mRNA sequence.

ACCESSION
CF743932
VERSION
CF743932.1 GI:37640271
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Mammalia; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1. (bases 1 to 801)

REFERENCE
NTH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-1@mail.nih.gov

AUTHORS
TITLE
JOURNAL
COMMENT
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefi.html>
This clone was contributed by the Brain Molecular Anatomy Project

[illegible]

with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lona-linker LL-sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

| | | | | | | | |
|-----------------------|--------|---|---------------|------------|-----|--------|------|
| Query Match | 28.8%; | Score | 737.6; | DB | 14; | Length | 779; |
| Best Local Similarity | 97.3%; | Pred | No. 5.3e-177; | | | | |
| Matches | 746; | Conservative | 1; | Mismatches | 20; | Indels | 0; |
| Gaps | 0; | | | | | | |
| QY | 262 | GAAGAAATATTTCTGGTCCATTGAATAGTGAAGCAAGTGTGTGCTGGATGATCA | 321 | | | | |
| DB | 12 | GAAGAAATATTTCTGGTCCATTGAATAGTGAAGCAAGTGTGTGCTGGATGATCA | 71 | | | | |
| QY | 322 | CTGCTATGACAGGACTGATGTGTGNGAAAAAGACAGCCCTGGAAGTGTGCTG | 381 | | | | |
| DB | 72 | CTGCTATGACAGGACTGATGTGTGNGAAAAAGACAGCCCTGGAAGTGTGCTG | 131 | | | | |
| QY | 382 | CTGTGAGGCAATATGTGTAAGTAAAGTCTCTTATTTCCGGAGATGGAAGTCACA | 441 | | | | |
| DB | 132 | CTGTGAGGCAATATGTGTAAGTAAAGTCTCTTATTTCCGGAGATGGAAGTCACA | 191 | | | | |
| QY | 442 | GCCCACTTCAATCTGTTTACACCGAGCCACCTATTACCAATCTGCTGATTCCT | 501 | | | | |
| DB | 192 | GCCCACTTCAATCTGTTTACACCGAGCCACCTATTACCAATCTGCTGATTCCT | 251 | | | | |
| QY | 502 | GGTACACATTAATGTTAAATTCAGGAATGTGATTTGCAATTTGGTGTACACATCA | 561 | | | | |
| DB | 252 | GGTACACATTAATGTTAAATTCAGGAATGTGATTTGCAATTTGGTGTACACATCA | 311 | | | | |
| QY | 562 | CAAGATGGCCCTACCTCTGTACTTGTCTCTACAGCCAGGACCCACCCACCTTC | 621 | | | | |
| DB | 312 | CAAGATGGCCCTACCTCTGTACTTGTCTCTACAGCCAGGACCCACCCACCTTC | 371 | | | | |
| QY | 622 | CCCATTTACTAGGTTGAGCCATTGAGCTGTTAGAAAGTGAAGCAAGGGAAGATTGG | 681 | | | | |
| DB | 372 | CCCATTTACTAGGTTGAGCCATTGAGCTGTTAGAAAGTGAAGCAAGGGAAGATTGG | 431 | | | | |
| QY | 682 | TTGTGCTTGAAAGCCAGTTGCTCAATGAATATGTGGCTGTCAAAATATTTCCAAATCA | 741 | | | | |
| DB | 432 | TTGTGCTTGAAAGCCAGTTGCTCAATGAATATGTGGCTGTCAAAATATTTCCAAATCA | 491 | | | | |
| QY | 742 | GGACAAACAGTCTCGGAGAAATGAATATGAAGTCTATAGTCTACCTGGAATGAACATCA | 801 | | | | |
| DB | 492 | GGACAAACAGTCTCGGAGAAATGAATATGAAGTCTATAGTCTACCTGGAATGAACATCA | 551 | | | | |
| QY | 802 | GAACATCTACAGTTCAATTTGTGAGAAAGAGGACCCAGTGTGGATGTGACCTGTG | 861 | | | | |
| DB | 552 | GAACATCTACAGTTCAATTTGTGAGAAAGAGGACCCAGTGTGGATGTGACCTGTG | 611 | | | | |
| QY | 862 | GCTAATCACAGCTTTCATGAAAGGGGTCTACGTGACAGCTTCTTAAAGGCTAATGTGT | 921 | | | | |
| DB | 612 | GCTAATCACAGCTTTCATGAAAGGGGTCTACGTGACAGCTTCTTAAAGGCTAATGTGT | 671 | | | | |
| QY | 922 | CTCTTGGAAATCACTTTTGTATTTGAGAAACCATGGCTAGAGGATGGGCAATTTTACA | 981 | | | | |
| DB | 672 | CTCTTGGAAATCACTTTTGTATTTGAGAAACCATGGCTAGAGGATGGGCAATTTTACA | 731 | | | | |
| QY | 982 | TGAGGATATACCTGGCTTAAAGATGGCCACAGGCTGCAATCTCTC | 1028 | | | | |

732 TGANGATATACCTGGCTTANAGATGNNACAGCCTGCACTCTCAC 778

RESULT 7

BQ960311

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ960311 899 bp mRNA linear EST 21-AUG-2002

AGENCOURT 8924486 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6466795

5', mRNA sequence.

BQ960311 GI:22375789

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 899)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM13991 row: k column: 20

High quality sequence stop: 650.

Location/Qualifiers

1. 899

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:6466795"

/tissue_type="retina"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_94"

/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 3.3 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

1;

Mismatches

5;

Indels

3;

 Gaps | 2; |

664 AGCAAGGGAAGATTGGTGTCTCTGGAAGCCAGTTCCTCAATGAATATGTGGCTGT 723

1 AGCAAGGGAAGATTGGTGTCTCTGGAAGCCAGTTCCTCAATGAATATGTGGCTGT 60

724 CAAATATTTCCAAATACAGCAACAAAGTCTGCGAGAAATGAATATGAAGTCTATAGTCT 783

61 CAAATATTTCCAAATACAGCAACAAAGTCTGCGAGAAATGAATATGAAGTCTATAGTCT 120

784 ACCTGGAATGAACATGAGAACATCTACAGTTCATTGGTGCAGAGAAAGAGGCCACAG 843

121 ACCTGGAATGAACATGAGAACATCTACAGTTCATTGGTGCAGAGAAAGAGGCCACAG 180

844 TGTGGATGTGACCTGTGGCTTAATCAGACATTTTCATGAAAAGGGCTCACTGTGACACTT 903

181 TGTGGATGTGACCTGTGGCTTAATCAGACATTTTCATGAAAAGGGCTCACTGTGACACTT 240

904 TCTTAAGGCTAATGTGTCTCTTGGAAATCACTTTGTTCATATTGAGAAACCATGGCTAG 963

241 TCTTAAGGCTAATGTGTCTCTTGGAAATCACTTTGTTCATATTGAGAAACCATGGCTAG 300

964 AGGATTGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAGCCTCAAT 1023

301 AGGATTGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAGCCTCAAT 360

1024 CTCTCAGGACATCAAAAGTAAATGCTGCTTGAAGAAACAACTCTGACAGCTTGCAAT 1083
b |||||
361 CTCTCAGGACATCAAAAGTAAATGCTGCTTGAAGAAACAACTCTGACAGCTTGCAAT 420
y |||||
1084 TGCTGACTTTGGTGGCTTAAAGTTCGAGGCTGCAAGTCTCAGGTGACACCAATGG 1143
b |||||
421 TGCTGACTTTGGTGGCTTAAAGTTCGAGGCTGCAAGTCTCAGGTGACACCAATGG 480
y |||||
1144 GCAGGTGGTACCCGAGGTATATGCTCCAGAGGTGTGGAGGGTGTCTATAAACTTCCA 1203
b |||||
481 GCAGGTGGTACCCGAGGTATATGCTCCAGAGGTGTGGAGGGTGTCTATAAACTTCCA 540
y |||||
1204 AAGGACGCAATTTCTGAGGATAGATATGTCAGCCATGGATGCTCTATGGGAATTCGC 1263
b |||||
541 AAGGACGCAATTTCTGAGGATAGATATGTCAGCCATGGATGCTCTATGGGAATTCGC 600
y |||||
1264 TTCTCGTTGCACTGCTGCAGATGACCCGCTAGATGAGTACATGTTACCATTTGAGGAGA 1323
b |||||
601 TTCTCGTTGCACTGCTGCAGATGACCCGCTAGATGAGTACATGTTACCATTTGAGGAGA 660
y |||||
1324 AATGGCCAGCATCCATCTCTTGAAGATATGCAAGAGTGTGTGTCATAAAAAAAGAG 1383
b |||||
661 AATGGCCAGCATCCATCTCTTGAAGATATGCAAGAGTGTGTGTCATAAAAAAAGAG 720
y |||||
1384 GCCTG-TTTTAAGAGATATTGG--CAGAAATGCGAGAA 1421
b |||||
721 GCCTGTTTAAAGATATTGGCCAGAAATGCGAGAA 761

BI827084 849 bp mRNA linear EST 04-OCT-2001
603075782F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167461 5',
mRNA sequence.
BI827084
BI827084.1 GI:15938646
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 849)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbe-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LIA411416 row: d column: 22
High quality sequence stop: 837.
Location/Qualifiers
1. 849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5167461"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dr primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber

FEATURES
source

ORIGIN
Query Match 28.1%; Score 721.2; DB 12; Length 849;
Best Local Similarity 92.6%; Pred. No. 8.3e-173;
Matches 789; Conservative 1; Mismatches 59; Indels 3; Gaps 3;
QY 690 GGAAAGCCAGGTCCTCAATGAATATGCTGCTGCAAAATATTTCCAAATACAGGACAAAC 749
Db 1 GGAAAGCCAGGTCCTCAATGAATATGCTGCTGCAAAATATTTCCAAATACAGGACAAAC 60
QY 750 AGTCCTGGCAGAAATGAATATGAAGTCTATAGTCTACCTGGAAATGAAGCATGAGAAACATAC 809
Db 61 AGTCATGGCCAAATGAATATGAGTCTACAGT-TGCCCTGGATGAGCATGAGAAACATAT 119
QY 810 TACAGTTCTATTTGCTGCAGAGAAAGAGCCACCTGCTGAGATGTGGACCTGTGGCTTAATCA 869
Db 120 TACAGTTCA-TGGTGCAGAAAAACAGGACCACTGTTGATGTGGATCTTTGGCTGATCA 178
QY 870 CAGCATTTCAATGAAAAAGGGCTCTACTGTGACAGTCTTTCTTAAGGCTTAATGTGCTCTTGA 929
Db 179 CAGCATTTCAATGAAAAAGGGTTCATATCAGACTTTCTTAAGGCTTAATGTGCTCTTGA 238
QY 930 ATCARTTTTCTATTTGTCAGAAACCATGGCTAGAGGATTTGSCATATTTACATGAGGATA 989
Db 239 ATGAACCTGTGTCTATTTGTCAGAAACCATGGCTAGAGGATTTGSCATATTTACATGAGGATA 298
QY 990 TACCTGGCTTTAAAGATGGCCACAAAGCTCTCAAGTCTCTCACAGGACATCAAAAGTAAAC 1049
Db 299 TACCTGGCTTTAAAGATGGCCACAAACCTGCCATATCTCACAGGACATCAAAAGTAAAC 358
QY 1050 ATGTGCTGTTGAAAAACAATCTGACAGCTTGCATTTGCTGACCTTTGGGTTGGCCTTAAAGT 1109
Db 359 ATGTGCTGTTGAAAAACAACCTGACAGCTTGCATTTGCTGACCTTTGGGTTGGCCTTAAAGT 418
QY 1110 TCGAGGCTGGCAAGTCTGCAAGTGCACCCATGGCAGGTTGGTACCCGAGGATATATGG 1169
Db 419 TTGAGGCTGGCAAGTCTGCAAGTGCACCCATGGCAGGTTGGTACCCGAGGATATATGG 478
QY 1170 CTCCAGAGGTTTGGAGGGTCTCTATAAATCTCCAAAGGACCATTTCTGAGGATAGATA 1229
Db 479 CTCCAGAGGTTTGGAGGGTCTATAAATCTCCAAAGGACCATTTCTGAGGATAGATA 538
QY 1230 TGTAAGCCATGGGATTTAGTCTCTATGGGAAATGGCTTCTGTTGCACTGCTGCAAGTGGAC 1289
Db 539 TGTAAGCCATGGGATTTAGTCTCTATGGGAAATGGCTTCTGTTGCACTGCTGCAAGTGGAC 598
QY 1290 CCGTAGATGAGTACATGTTTACCATTTGAGGAGAAATTTGGCCAGCATCCATCTCTTGAAG 1349
Db 599 CTGTAGATGAGTACATGTTTGCATTTGAGGAGAAATTTGGCCAGCATCCATCTCTTGAAG 658
QY 1350 ATATGCGAGGAGTGTGTCGATATAAAAAAGAGGCTGTTTAAAGAGATTATTTGGCAGA 1409
Db 659 ACATGCGAGGAGTGTGTCGATATAAAAAAGAGGCTGTTTAAAGAGATTATTTGGCAGA 717
QY 1410 AACATGCGAGGATGGCATGCTCTGTGAACCATAGAAATGTTGGGATCATGTATGATG 1469
Db 718 AACATGCGAGGATGGCATGCTCTGTGAACCATAGAAATGTTGGGATCATGTATGATG 777
QY 1470 AAGCCAGGTTATCAGCTGGATGTAGGTGAAGAAATTTACTCAGATGCCAAAGACTTAACA 1529
Db 778 AAGCCAGGTTATCAGCTGGATGTAGGTGAAGAAATTTACTCAGATGCCAAAGACTTAACA 837
QY 1530 ATATCATTACTA 1541
Db 838 ATATCATTACTA 849

RESULT 9
CB245658
LOCUS
DEFINITION UI-M-FY0-cdt-n-09-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
717 bp mRNA linear EST 09-JUL-2003
CB245658
UI-M-FY0-cdt-n-09-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone

IMAGE:6834346 5', mRNA sequence.
CB245658
VERSION CB245658.1 GI:28367302
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 717)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. 717
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6834346"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_F10"
/notes="Organ: Brain; Vector: pYX-Asc; Site: 1; Ecor I; Site 2: Not I; The library was constructed according to Bonafo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., Program coordinator."

ORIGIN
Query Match 27.5%; Score 705; DB 14; Length 717;
Best Local Similarity 99.7%; Pred. No. 1.1e-168;
Matches 716; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1236 CCATGGATTAGTCTATGGAAATGGCTTCTGTTGCACTGCTGAGATGACCGTAG 1295
1 CCATGGATTAGTCTATGGAAATGGCTTCTGTTGCACTGCTGAGATGACCGTAG 60
QY 1296 ATGAGTACATGTACCATTTGAGGAGAAATGGCCAGCATCCATCTTTGAAGATATGC 1355
61 ATGAGTACATGTACCATTTGAGGAGAAATGGCCAGCATCCATCTTTGAAGATATGC 120
QY 1356 AGGAAGTGTGTGCATAAAAAAGAGCGCCGTTTAAAGATATTTCGCGAACAATG 1415
121 AGGAAGTGTGTGCATAAAAAAGAGCGCCGTTTAAAGATATTTCGCGAACAATG 180
QY 1416 CAGGAATGGCAATGCTCTGTGAAACGATAGAGAATGTTGGATCATGATGAGAGGCCA 1475
181 CAGGAATGGCAATGCTCTGTGAAACGATAGAGAATGTTGGATCATGATGAGAGGCCA 240
QY 1476 GGTTATACGCTGGATGTGTAGGTGAAGAATTAATCTCAGATGCAAGACTTAACAATATCA 1535

Db 241 GGTTATACGCTGGATGTGTAGGTGAAGAATTAATCTCAGATGCAAGACTTAACAATATCA 300
QY 1536 TTACTACAGAGGACATTTGTACAGTGGTCACAATGGTCAACAATGTTGACTTTCTCTCCCA 1595
Db 301 TTACTACAGAGGACATTTGTACAGTGGTCACAATGGTCAACAATGTTGACTTTCTCTCCCA 360
QY 1596 AAGAACTTAGTCTATGATGGTGGACCGCTCTGTACACACTGAGGACTGGGACTCTGAAC 1655
Db 361 AAGAACTTAGTCTATGATGGTGGACCGCTCTGTACACACTGAGGACTGGGACTCTGAAC 420
QY 1656 GGAGCTGCTAAGCTAAGCAAGAGTCTTAGTCTGATTTCTGTGTGAATGATGATGATGCC 1715
Db 421 GGAGCTGCTAAGCTAAGCAAGAGTCTTAGTCTGATTTCTGTGTGAATGATGATGATGCC 480
QY 1716 TCCAGGACATGTACCAAGCAGCCCTTGTGAAAGCATGCTGCGAGATGATGATGATGCC 1775
Db 481 TCCAGGACATGTACCAAGCAGCCCTTGTGAAAGCATGCTGCGAGATGATGATGATGCC 540
QY 1776 GAACTTACTGATGCTGTCAGCAGCAGATATCAAGAGAGTCTTAAGGGGAAAAAGCTGCA 1835
Db 541 GAACTTACTGATGCTGTCAGCAGCAGATATCAAGAGAGTCTTAAGGGGAAAAAGCTGCA 600
QY 1836 ACTGTAAGAACTTCTGAAATGTTACTCGAAGAAATGTCGCCCTCTCCAAATCAAGGATCT 1895
Db 601 ACTGTAAGAACTTCTGAAATGTTACTCGAAGAAATGTCGCCCTCTCCAAATCAAGGATCT 660
QY 1896 TTTGACCTGGTAAATCAAGTATTTGCAAACTGCAACATGATGATGATGATGATGATGTC 1953
Db 661 TTTGACCTGGTAAATCAAGTATTTGCAAACTGCAACATGATGATGATGATGATGATGTC 717

RESULT 10
BU704233
LOCUS BU704233
DEFINITION UI-M-FOO-bzs-k-15-0-UI.r1 NIH-BMAP_F00 Mus musculus cDNA clone
IMAGE:6406790 5', mRNA sequence.
BU704233
VERSION BU704233.1 GI:23632150
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 728)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. 728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6406790"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_F00"
/notes="Organ: Brain; Vector: pYX-Asc; Site: 1; Ecor I; Site 2: Not I; The library was constructed according

RESULT 11

source

| | Query Match | 27.1%; | Score 595.8; | DB 14; | Length 742; |
|----|-----------------------|---|---------------------|-----------|-------------|
| | Best Local Similarity | 99.8%; | Pred. No. 2.5e-166; | | |
| | Matches 722; | Conservative 0; | Mismatches 7; | Indels 2; | Gaps 2 |
| Qy | 1560 | TGCTCACAATGGTGACAAAATGTCACATTTCTCTCCCAAGAATCTAGTCTATGATGGTGGC | 1611 | | |
| Db | 1 | TGCTCACAATGGTGACAAAATGTCACATTTCTCTCCCAAGAATCTAGTCTATGATGGTGGC | 60 | | |
| Qy | 1620 | ACCGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTTAAGCTAAGGAAAGTG | 1671 | | |
| Db | 61 | ACCGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTTAAGCTAAGGAAAGTG | 120 | | |
| Qy | 1680 | CTTAGTTGATTTTCTGTGTGAAATGAGTAGGATGCTCCAGGACATGTACGCAAGCAGCC | 173 | | |
| Db | 121 | CTTAGTTGATTTTCTGTGTGAAATGAGTAGGATGCTCCAGGACATGTACGCAAGCAGCC | 180 | | |
| Qy | 1740 | CTTTGTGGAAAGCATGGATCTGGGAGATGGATCTGGGAAACTTACTGCATCGTCTGCAGC | 179 | | |
| Db | 181 | CTTTGTGGAAAGCATGGATCTGGGAGATGGATCTGGGAAACTTACTGCATCGTCTGCAGC | 240 | | |
| Qy | 1800 | ACAGATATGAAGAGAGTCTTAAGGGAAAAGCTGCAAACTGTAAAGAACTTCTTGAAAATGT | 185 | | |

| | Query Match | 27.1%; | Score 595.8; | DB 14; | Length 742; |
|----|-----------------------|---|---------------------|-----------|-------------|
| | Best Local Similarity | 99.8%; | Pred. No. 2.5e-166; | | |
| | Matches 722; | Conservative 0; | Mismatches 7; | Indels 2; | Gaps 2 |
| Qy | 1560 | TGCTCACAATGGTGACAAAATGTCACATTTCTCTCCCAAGAATCTAGTCTATGATGGTGGC | 1611 | | |
| Db | 1 | TGCTCACAATGGTGACAAAATGTCACATTTCTCTCCCAAGAATCTAGTCTATGATGGTGGC | 60 | | |
| Qy | 1620 | ACCCTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTTAAGCTAAGGAAAGTG | 1671 | | |
| Db | 61 | ACCCTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTTAAGCTAAGGAAAGTG | 120 | | |
| Qy | 1680 | CTTAGTTGATTTTCTGTGTGAAATGAGTAGGATGCTCCAGGACATGTACGCAAGCAGCC | 173 | | |
| Db | 121 | CTTAGTTGATTTTCTGTGTGAAATGAGTAGGATGCTCCAGGACATGTACGCAAGCAGCC | 180 | | |
| Qy | 1740 | CTTTGTGGAAAGCATGGATCTGGGAGATGGATCTGGGAAAATTACTGCATCGTCTGCAGC | 179 | | |
| Db | 181 | CTTTGTGGAAAGCATGGATCTGGGAGATGGATCTGGGAAAATTACTGCATCGTCTGCAGC | 240 | | |
| Qy | 1800 | ACAGATATGAAGAGAGTCTAAGGGAAAAGCTGCAAACTGTAAAGCACTTCTTGAAAATGT | 185 | | |

241 ACAGATATGAAGAGAGTCTAAGGAAAGCTGCAAACTGTAAAGAACTTCTGAAAATGT 300
1860 ACTCGAAGAAATGTGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGATTT 1919
301 ACTCGAAGAAATGTGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGATTT 360
1920 TCGAAAACATGACATGAGATTTCTTAATGCTGTGACAGACACTAATTCCTTAAATGAAC 1979
361 TG-AAAACATGACATGAGATTTCTTAATGCTGTGACAGACACTAATTCCTTAAATGAAC 419
1980 TACTGCTATTTTTTTTAAATGAAGAACTTTTCAATTCAGATTTTAAAGGGTAACCTTT 2039
420 TACTGCTATTTTTTTTAAATGAAGAACTTTTCAATTCAGATTTTAAAGGGTAACCTTT 479
2040 TATTGCTATTTGCTGTTTCTAATAATGACATATGCTAATGCTAATGCTAATGCTAATGCT 2099
480 TATTGCTATTTGCTGTTTCTAATAATGACATATGCTAATGCTAATGCTAATGCTAATGCT 539
2100 GAATGTGTAGTGTGCTGCTGTTCTGTGTATCATAGTCATCAAAAGTGGGTACAGTAAAGAG 2159
540 GAATGTGTAGTGTGCTGCTGTTCTGTGTATCATAGTCATCAAAAGTGGGTACAGTAAAGAG 599
2160 GCTTCCAGCATTAATTTAACTCCCTCAACAGGATATACCTCAGTCCACGGTGTGTAA 2219
600 GCTTCCAGCATTAATTTAACTCCCTCAACAGGATATACCTCAGTCCACGGTGTGTAA 658
2220 ATTATAAATTTGAAACACTTAACAGATTTGAATAATCAGTCATGTTTATTAACAGG 2279
659 ATTATAAATTTGAAACACTTAACAGATTTGAATAATCAGTCATGTTTATTAACAGG 718
2280 TTAATTACAAA 2290
719 TAAAAAATAA 729

RESULT 12
CF271903
LOCUS
DEFINITION
IMAGE:15197010 NIH_MGC_195 Homo sapiens cDNA clone
IMAGE:7002205 5', mRNA sequence.

CF271903
CF271903.2 GI:38558822
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 763)
NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
On Aug 12, 2003 this sequence version replaced gi:33627815.
Contact: Daniela S. Gerhard, ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: IRBK4 row: g column: 11
High quality sequence stop: 706.

FEATURES
source
1. 763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002205"
/tissue_type="mixed"

/lab_host="DH5A (T1 phage-resistant)"
/clone_lib="NIH_MGC_195"
/note=vector: pDNR-Dual; Site 1: loxP-Sall; Site 2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxP sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK_presV.dat
a Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.1%; Score 668.8; DB 14; Length 763;
Best Local Similarity 94.4%; Pred. No. 1.9e-159;
Matches 591; Conservative 1; Mismatches 40; Indels 0; Gaps 0;
QY 1 CTCGAGGAAGACCCAGGAACTGGATATCTAGCAGAACTTCTAGCGGCTTCTCCGCG 60
DB 32 CTCCGAGGAAGACCCAGGAACTGGATATCTAGCAGAACTTCTCCGGAATCCCGCG 91
QY 61 CCTCGGAATGAGGAGCTGCTGCAAGTTGGGCTTGGCGCTTCTTATCTCTGCTC 120
DB 92 CCTCGGAATGAGGAGCTGCTGCAAGTTGGGCTTGGCGCTTCTTATCTCTGCTC 151
QY 121 TTCAGTGTCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTTCTTTTAAATGCTA 180
DB 152 TTCAGTGTCTATCTTGGTAGATCAGAACTCAGGAGTGTCTTTTCTTTAATGCTA 211
QY 181 GGAAAGAGCAGAACCAACCACTGGTGTGAACCTTGTGATGATGATGATGATGATGAT 240
DB 212 GGAAAGAGCAGAACCAACCACTGGTGTGAACCTTGTGATGATGATGATGATGATGAT 271
QY 241 GCGACATTTGTTTGTACCTCGGAGAAATATTTCTGTTTCCATTGAAATAGTGAAGCA 300
DB 272 GCGGCATTTGTTTGTACCTCGGAGAAATATTTCTGTTTCCATTGAAATAGTGAAGCA 331
QY 301 TTGTTGGCTGGATGATATCAACTGTATGACAGACTGATTGTGTGAAAGAAAGACAG 360
DB 332 TTGTTGGCTGGATGATATCAACTGTATGACAGACTGATTGTGTGAAAGAAAGACAG 391
QY 361 CCTGGAAGTGTACTTTTGTGCTGTGAGGCAATATGTGTAATGAAAGTTCTCTATT 420
DB 392 CCTGGAAGTGTACTTTTGTGCTGTGAGGCAATATGTGTAATGAAAGTTCTCTATT 451
QY 421 TCCGGAGATGGAAGTCAACAGCCCACTTCAAAATCCAGTTACACCTAGCCACCCTAT 480
DB 452 TCCGGAGATGGAAGTCAACAGCCCACTTCAAAATCCAGTTACACCTAGCCACCCTAT 511
QY 481 CAACATCTGCTGTATCTTCTGTTGACCACTTAATGTTAATGACAGAAATGTCATTG 540
DB 512 CAACATCTGCTGTATCTTCTGTTGACCACTTAATGTTAATGACAGAAATGTCATTG 571
QY 541 ATTTTGGGTGACAGATCAGATCAGATGAGGCTACCTCTCTCTCTCTCTCTCTCTCA 600
DB 572 ATTTTGGGTGACAGATCAGATCAGATGAGGCTACCTCTCTCTCTCTCTCTCTCTCA 631
QY 601 CCCAGGACCAACCCCACTTCCCCTTCTAGGTTGAGGCTTGAAGCCATTCGAGCTTTAGA 660
DB 632 CCCAGGACCAACCCCACTTCCCCTTCTAGGTTGAGGCTTGAAGCCATTCGAGCTTTAGA 691
QY 661 GAAAGCAAGGGGAGATTTGTTGCTGTAAGCCAGCTTGTCTCAATGATATATGCTGC 720
DB 692 GAAAGCAAGGGGAGATTTGTTGCTGTAAGCCAGCTTGTCTCAATGATATATGCTGC 751

[illegible]

124 GCTACGCTCTCGCGCAACAGCTCTGCGCACCATCGAGCTCGTGAAGAAGGCTGCTGGC 183
1309 TGGATGATCACTGCTATGACAGGACTGATTGTGTGAAAAAAGACACGCCCTGAG 368
184 TAGATGACTTCAACTGCTACGATAGGACAGGAGTGTGCGCCACTGAGGAGAACCCCGAG 243
369 TGTACTTTTGTGCTGAGGGCAATATGTCTAATGAAAAAGTTCTCTTATTTTCGGAGA 428
244 TGTACTTCTGCTGTGAAGGCAACTCTGCAACGAGGCTTCACTCATTTTGCAGAGG 303
429 TGGAAGTACACAGACCCACTTCAATCTCTGTATACACCGAAGCCACCTATTACAACATTC 488
304 CTGGGGCCCGGAAGTCAAGTACAGGACCAACCCCGACAGCCGCCCTGCTCAAGGTGC 363
489 TGTCTATTCTTGGTACCACTAATGTAAATGACAGGAATGTGTCAATTTGTGATTTGGG 548
364 TGGCCTACTCACTGCTGCCATCGGGGGCTTTCCCTCATCTGCTCTGCTGGGCTTTTGA 423
549 TGTACAGACATCAAGATGCCCTACCTCTCTGTACTTGTTCCTACTCAAGACCCAGGAC 608
424 TGTACCGGCATCGACGCCCTTACGGTCAATGTGGACATCC---ATGAGGACCCCTGGGC 480
609 CACCCCCACCTTCCCATTAATAGGTTGAAGCCATTCAGCTGTAGAAAGTGAAGCAA 668
481 CTCACACCATCCCTCTGCTGGCCCTGAAGCCACTGCAGCTGTGGAGATCAAGSCTC 540
669 GGGGAAGATTGTTGTCTGGAAGCCCAAGTGTCTCAATGAATATGTGGCTGTCAAAA 728
541 GGGGGCTTTGGTGTCTGGAAGGCCCAAGCTCATGAATGACTTTGTAGCTGTCAAGA 600
729 TATTTCCAAATACAGGACAAACAGCTCTGGCAGATGAATATGAAGTCTATAGTCTACCTG 788
601 TCTTCCCACTCCAGGACAAAGCAGTGTGCGCAGTGAACGGGAGATCTTCAGCACACCTG 660
789 GAATGAGCATGAGAACATATCACTGATTTCTGTTGCGAGAGAAAGAGGACCAAGTGTG 848
661 GCATGAGCAGCAGAACCTGCTTACAGTTTCAATTTGCTGCGAGAGAGGAGCTCCAACTCG 720
849 ATGTGGACCTGTGGCTAATCAAGCATTTCTATGAAGAGGCTCACTGTGACAGCTTTCTTA 908
721 AAGTAGAGCTGTGGCTCATACGGCTTCCATGACAGGGCTCCCTCAGGATTAATCTCA 780
909 AGGCTAATGTGGTCTCTTGGAAATCACTTTGTCTATTTCAGAAACCATGGCTAGAGAT 968
781 AGGGGAACATCATACATGGAACGAAGTGTGTATGTAGCAGAGACGATGTACAGAGGCC 840
969 TGGCATATTTACATGAGGATATACC---TGGCTTAAAGATGCGCACAGCCCTGCAATCT 1025
841 TCTCATACCTGATGAGGATGTGCCCTGCTGCGGAGGGGCCACAAGCCGCTCTATTG 900
1026 CTCACAGGACATCAAAAGTAAATATGCTGTGAAGAACATCTGACAGCTTGCATTG 1085
901 CCCACAGGACTTTAAAGTAAAGATGATTTGCTGAAGAGCGACCTCAGAGCCGTGCTGG 960
1086 CTGACTTTGGGTTGGCTTAAAGTTTCAGGCTGGCAAGTCTGAGGTGACACCCATGGGC 1145
961 CTGACTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1020
1146 AGGTTGGTACCCGGAGGTATATGGCTCCAGAGGTGTGGAGGGTGTCTATAAACTTCCAAA 1205
1021 AGGTAGGCACGAGACGGTACATGGCTCTCTGAGTGTCTGAGGGAGCCATCAACTCCAGA 1080
1206 GGGAGCGATTTCTGAGGATAGATATGTACGCCATGGGATTAGTCTTATGGGAATGGCTT 1265
1081 GAGATGCCCTTCTGCGCATTTGACATGTATGCCATGGGGTTGGTGTGTGGAGGCTTGTGT 1140
1266 CTGTTGCACTGCTCCAGTGGACCGGTAGATGATGATGTTTACCATTTGAGGAAGAAA 1325
1141 CTCGCTGAAGGCTGACAGCCCGTGGATGAGTACATGCTGCCCCCTTTGAGGAAGAGA 1200
1326 TTGGCCAGCATCCATCTCTTTGAAGATATGAGGAAGTTGTTGTGCAATAAAAAAAGAGGC 1385
1201 TTGGCCAGCACCTTCTGTTGGAGAGCTGACAGAGGTTGTTGTTGTCACAGAGATGAGGC 1260

QY 1386 CTGTTTTAAGAGATTATTGGCAGAAACATGAGGAATGGCAATGCTCTGTGAACGATAG 1445
Db 1261 CCACCATTAAGATCACTGGTTGAACACCCCGGCTGGCCAGCTTTGTGTGACCATCG 1320
QY 1446 AAGAAATGTTGGGATCATGATCGAAGCCAGGTTATCAGCTGGATGTGTAGGTGAAGAA 1505
Db 1321 AGGAGTGTGGGACCATGATGACAGAGGCTGGCTTGTCCGCGGCTGTGTGGAGAGCGGG 1380
QY 1506 TTACTCAGATCCAAAGACTAAACAATATCATTTACTACAGAGGACATTTGTAACAGTGTCA 1565
Db 1381 TGTCCCTGATTTCGGAGGTTCGATGATGACACTTACCTCGGACTGTCTCGTTTCCCTGGTGA 1440
QY 1566 CAATGGTGACAAATGTTGACCTTTCCTCCCAAGAAATCTAGTCTATGA 1612
Db 1441 CCTCTGTACCAATGTGGACCTGCCCCCTTAAAGAGTCAAGCATCTAA 1487

Search completed: February 26, 2004, 07:23:33
Job time : 5898 secs

| | | | | | | |
|----|--------|------|-------|---|----------|-----------|
| 1 | 2561.2 | 99.9 | 2563 | 2 | AAx22685 | Human act |
| 2 | 2561.2 | 99.9 | 2563 | 4 | AAc85299 | Human act |
| 3 | 2561.2 | 99.9 | 2563 | 6 | AAK98725 | cDNA enco |
| 4 | 2566.4 | 99.7 | 2563 | 2 | AAQ31912 | Human act |
| 5 | 2566.4 | 99.7 | 2563 | 2 | AAx22683 | Mouse act |
| 6 | 2566.4 | 99.7 | 2563 | 4 | AAx22683 | Mouse act |
| 7 | 2566.4 | 99.7 | 2563 | 6 | AAK85297 | Mouse act |
| 8 | 2551.6 | 99.6 | 2563 | 6 | AAK98720 | cDNA enco |
| 9 | 2080.4 | 81.2 | 2277 | 9 | AAQ31910 | Mouse act |
| 10 | 2057.4 | 80.3 | 2122 | 2 | AAK53787 | Primary r |
| 11 | 1987.6 | 77.5 | 2951 | 5 | AAb24517 | Mouse act |
| 12 | 1378.2 | 77.2 | 5286 | 6 | ABv25076 | Human pro |
| 13 | 1903 | 74.2 | 2382 | 6 | AAc94867 | Human DNA |
| 14 | 1440 | 56.2 | 1442 | 4 | ABz35028 | Human gen |
| 15 | 1406 | 54.9 | 1466 | 3 | AAh20184 | Mouse two |
| 16 | 839.6 | 32.8 | 1024 | 7 | AAa55943 | Polynucle |
| 17 | 697.2 | 27.2 | 6255 | 4 | ABz83451 | Toxicolog |
| 18 | 697.2 | 27.2 | 10409 | 4 | AAk89767 | Human dig |
| 19 | 678.2 | 26.5 | 2335 | 2 | AAk89769 | Human dig |
| 20 | 678.2 | 26.5 | 2335 | 4 | AAx22684 | Xenopus a |
| 21 | 678.2 | 26.5 | 2335 | 6 | AAc85298 | Xenopus a |
| 22 | 676.6 | 26.4 | 2335 | 2 | AAK98721 | cDNA enco |
| 23 | 627.2 | 24.5 | 1651 | 2 | AAQ31911 | Xenopus a |
| 24 | 627.2 | 24.5 | 1651 | 2 | AAV71967 | Mouse Act |

PS Disclosure

Disclosure: Page: 28pp: English.

Disclosure: Page: 28pp; English.

This sequence encodes a human activin receptor polypeptide. The nucleic acid molecules of the invention are useful as probes for the identification of additional members of the activin/TGF-superfamily of receptor proteins, and the coding sequences can be used for the recombinant expression of the receptor proteins or functional fragments of them. They may also be used to study the function and activity of activin receptor polypeptides in cells and to identify agents which will modulate activin receptor expression and activity for use in treating conditions such as carcinogenesis, wound healing, disorders of the immune or central nervous systems and especially the reproductive system (where they may be used to control fertility in humans, domestic and commercial animals). NOTE: This sequence does not appear in the specification but has been created from the mouse activin receptor sequence (represented in AAX22683) as described in the specification

| | | | | |
|-----------------------|--------------|--------------|------------|-------------|
| Query Match | 99.9% | Score 2561.2 | DB 2 | Length 2563 |
| Best Local Similarity | 100.0% | Pred. No. 0 | | |
| Matches 2563 | Conservative | 0 | Mismatches | 0 |
| | | | Indels | 0 |
| | | | Gaps | 0 |

| | | | | | | |
|---|------------------------|---------------------------|--------------------------|--------------------|--------|----|
| 1 | CTCCGAGGAAGACCCAGGGAAC | TGGATATCTACGAGAACTTC | CCTACGGGCTTC | CCGGCG | 60 | |
| | 1 | CTCCGAGGAAGACCCAGGGAAC | TGGATATCTACGAGAACTTC | CCTACGGGCTTC | CCGGCG | 60 |
| | 61 | CTCCGGGAAAAATGGGAGCTGCT | GCAAAAGTTGGGTTGCGCGTCT | TTCTTATCTCTTCGTC | 120 | |
| | 61 | CTCCGGGAAAAATGGGAGCTGCT | GCAAAAGTTGGGTTGCGCGTCT | TTCTTATCTCTTCGTC | 120 | |
| | 121 | TTCCAGTGTCTATACTTTGGCAGAT | CAGAAACTCAGAGTGTCTTTCT | TTTAATGCTAATTTG | 180 | |
| | 121 | TTCCAGTGTCTATACTTTGGCAGAT | CAGAAACTCAGAGTGTCTTTCT | TTTAATGCTAATTTG | 180 | |
| | 181 | GGAAAARGACAGAAACCAACAGAC | TGCTGTTGAACTTTGCTATGAT | GGTATAAAGATAAACG | 240 | |
| | 181 | GGAAAARGACAGAAACCAACAGAC | TGCTGTTGAACTTTGCTATGAT | GGTATAAAGATAAACG | 240 | |
| | 241 | GCACATTTGTTTGTCTACCTGGAA | GAATATTTCTGGTTCCATTGAAAT | TAGTGAAGCAAGG | 300 | |
| | 241 | GCACATTTGTTTGTCTACCTGGAA | GAATATTTCTGGTTCCATTGAAAT | TAGTGAAGCAAGG | 300 | |
| | 301 | TTGTTGGCTGGATGATATCAACTG | CTATGACAGACATGATTTGCTG | TGNGAAAAAAGACAG | 360 | |
| | 301 | TTGTTGGCTGGATGATATCAACTG | CTATGACAGACATGATTTGCTG | TGNGAAAAAAGACAG | 360 | |
| | 361 | CCCTGAAGTGATCTTTGTTGCTGT | GAGGCAATATGTGTAAATGAA | AGTTCTCTTATTT | 420 | |
| | 361 | CCCTGAAGTGATCTTTGTTGCTGT | GAGGCAATATGTGTAAATGAA | AGTTCTCTTATTT | 420 | |
| | 421 | TCCGGAGATGGAAGTCAACAGCC | CACTTCAAACTCTGTTACACG | GAAGCCACCTTATTA | 480 | |
| | 421 | TCCGGAGATGGAAGTCAACAGCC | CACTTCAAACTCTGTTACACG | GAAGCCACCTTATTA | 480 | |
| | 481 | CAACATTCCTGCTGATTCCTTTGG | TACCACTAATGTTAAATTGCA | GAAATTTGTCATTTGTC | 540 | |
| | 481 | CAACATTCCTGCTGATTCCTTTGG | TACCACTAATGTTAAATTGCA | GAAATTTGTCATTTGTC | 540 | |
| | 541 | ATTTTGGGTGATACAGACATCA | AGATGCCCTACCTCTGTA | CTTGTTCCTACTCAAGA | 600 | |
| | 541 | ATTTTGGGTGATACAGACATCA | AGATGCCCTACCTCTGTA | CTTGTTCCTACTCAAGA | 600 | |
| | 601 | CCCAGGACCCACCCACCTTCCC | CAATTAAGGGTTGAAGCCATT | TCAGCTGTTAGAAGT | 660 | |
| | 601 | CCCAGGACCCACCCACCTTCCC | CAATTAAGGGTTGAAGCCATT | TCAGCTGTTAGAAGT | 660 | |
| | 661 | GAAAGCAAGGGGAAGATTTGGT | TGTCTGGAAGCCAGTTGCTCA | ATGAAATATGTCGC | 720 | |
| | 661 | GAAAGCAAGGGGAAGATTTGGT | TGTCTGGAAGCCAGTTGCTCA | ATGAAATATGTCGC | 720 | |
| | 721 | TGTCAAAATATTTCCNATACAG | CAAAACAGTCCTCGCGAAT | TGAATATGAAGTCTATAG | 780 | |
| | 721 | TGTCAAAATATTTCCNATACAG | CAAAACAGTCCTCGCGAAT | TGAATATGAAGTCTATAG | 780 | |

| | | | | | | | | | | | | | | | |
|----|------|-------|-----------|-------------------|------------|------------|----------|----------|-----------|---------|----------|------------|---------|-------|------|
| Qy | 781 | TC | TACCTGGAA | TGAAGCATGAGAA | CATAC | TACAGTTCA | TGGTG | CAGAGAA | AGAGGCAC | 840 | | | | | |
| Db | 781 | TC | TACCTGGAA | TGAAGCATGAGAA | CATAC | TACAGTTCA | TGGTG | CAGAGAA | AGAGGCAC | 840 | | | | | |
| Qy | 841 | CAG | TGTGGA | TGTGACCTGTGGCTAA | TAC | CAGCATTT | CATGAAA | AGGGCT | CAC | TGTCAGA | 900 | | | | |
| Db | 841 | CAG | TGTGGA | TGTGACCTGTGGCTAA | TAC | CAGCATTT | CATGAAA | AGGGCT | CAC | TGTCAGA | 900 | | | | |
| Qy | 901 | CTT | TCTTAA | GGCTAATGTGGTCTCTT | TGAA | TCA | CTTTGT | TCATAT | TG | CAGAAA | CCATGGC | 960 | | | |
| Db | 901 | CTT | TCTTAA | GGCTAATGTGGTCTCTT | TGAA | TCA | CTTTGT | TCATAT | TG | CAGAAA | CCATGGC | 960 | | | |
| Qy | 961 | TAG | AGGAT | TGGCATA | TTTAT | TACAT | GAGGATAT | ACTCG | GCTTAA | AGAT | TGGCC | CACAGCCTGC | 1020 | | |
| Db | 961 | TAG | AGGAT | TGGCATA | TTTAT | TACAT | GAGGATAT | ACTCG | GCTTAA | AGAT | TGGCC | CACAGCCTGC | 1020 | | |
| Qy | 1021 | AAT | CTCTC | CACAGGACATCA | AAAGTAA | AAATCT | GCTGTTG | AA | AAAA | CAATCT | GT | CACAGCTTG | 1080 | | |
| Db | 1021 | AAT | CTCTC | CACAGGACATCA | AAAGTAA | AAATCT | GCTGTTG | AA | AAAA | CAATCT | GT | CACAGCTTG | 1080 | | |
| Qy | 1081 | CATT | GCTC | GACTTTTG | GGTGGCC | TTAA | AGTTG | CAGGCTGG | CAAGTCT | GT | CAGGTG | CACACCCA | 1140 | | |
| Db | 1081 | CATT | GCTC | GACTTTTG | GGCCTTAA | AGTTG | CAGGCTGG | CAAGTCT | GT | CAGGTG | CACACCCA | 1140 | | | |
| Qy | 1141 | TGG | CAGGTTGG | TATACCCG | GAGGTATATG | CTCC | CAGAGTGT | TGG | AGGGTGT | TAA | AACTT | 1200 | | | |
| Db | 1141 | TGG | CAGGTTGG | TATACCCG | GAGGTATATG | CTCC | CAGAGTGT | TGG | AGGGTGT | TAA | AACTT | 1200 | | | |
| Qy | 1201 | CCAA | AGGACG | CAATTTCT | GAGGATAG | ATATG | TACGCCAT | GGGATT | TACT | CTCTAT | CGGAA | TT | 1260 | | |
| Db | 1201 | CCAA | AGGACG | CAATTTCT | GAGGATAG | ATATG | TACGCCAT | GGGATT | TACT | CTCTAT | CGGAA | TT | 1260 | | |
| Qy | 1261 | GGCT | TCTG | TCTGCAC | TGCTGC | AGATG | CA | CCG | TAGATG | AGTAC | ATGTT | TAGGA | 1320 | | |
| Db | 1261 | GGCT | TCTG | TCTGCAC | TGCTGC | AGATG | CA | CCG | TAGATG | AGTAC | ATGTT | TAGGA | 1320 | | |
| Qy | 1321 | AGAA | ATTG | CGCAG | CATCC | ATCTCTT | GAA | GATAT | CG | AGAA | GT | TGTC | ATA | 1380 | |
| Db | 1321 | AGAA | ATTG | CGCAG | CATCC | ATCTCTT | GAA | GATAT | CG | AGAA | GT | TGTC | ATA | 1380 | |
| Qy | 1381 | GAG | GCCTG | TTTAA | GAGATTTAT | TGG | C | GAA | CA | ATG | CAGGAA | TG | CAATG | CTGTG | 1440 |
| Db | 1381 | GAG | GCCTG | TTTAA | GAGATTTAT | TGG | C | GAA | CA | ATG | CAGGAA | TG | CAATG | CTGTG | 1440 |
| Qy | 1441 | GAT | AGAA | TATG | TGGGATCA | TGATG | CAGAA | GCCAG | CA | GTG | TAT | CAGTGT | AGTGA | 1500 | |
| Db | 1441 | GAT | AGAA | TATG | TGGGATCA | TGATG | CAGAA | GCCAG | CA | GTG | TAT | CAGTGT | AGTGA | 1500 | |
| Qy | 1501 | AAGAA | TTTACT | TCAGATG | CA | AAAGACTTAA | CAAA | TAT | CATTTACTA | CAGAGG | CA | TTGTA | ACAGT | 1560 | |
| Db | 1501 | AAGAA | TTTACT | TCAGATG | CA | AAAGACTTAA | CAAA | TAT | CATTTACTA | CAGAGG | CA | TTGTA | ACAGT | 1560 | |
| Qy | 1561 | GGT | CACA | ATGGTGA | CAAA | TGTTG | ATCTTCT | CCCAA | AGAA | TCTAGT | CTAT | GATG | TGTC | 1620 | |
| Db | 1561 | GGT | CACA | ATGGTGA | CAAA | TGTTG | ATCTTCT | CCCAA | AGAA | TCTAGT | CTAT | GATG | TGTC | 1620 | |
| Qy | 1621 | CCG | TCTGT | CA | CAC | TGAGGACTGG | CA | CTCTG | AA | CTG | AGAGTCT | AGTCT | AGTCT | 1680 | |
| Db | 1621 | CCG | TCTGT | CA | CAC | TGAGGACTGG | CA | CTCTG | AA | CTG | AGAGTCT | AGTCT | AGTCT | 1680 | |
| Qy | 1681 | TTAG | TTGATTTT | CTGTGT | GAAATG | AGTATG | AGGATG | CC | TCC | AGGACAT | GTAC | G | CAAGCAG | 1740 | |
| Db | 1681 | TTAG | TTGATTTT | CTGTGT | GAAATG | AGTATG | AGGATG | CC | TCC | AGGACAT | GTAC | G | CAAGCAG | 1740 | |
| Qy | 1741 | CTT | TG | GAAG | ACATG | ATG | ATG | ATG | ATG | ATG | ATG | ATG | ATG | 1800 | |
| Db | 1741 | CTT | TG | GAAG | ACATG | ATG | ATG | ATG | ATG | ATG | ATG | ATG | ATG | 1800 | |
| Qy | 1801 | CAG | ATATGA | AGAG | AGTCTA | AGG | GAAG | AGCTG | CAAA | CTG | TA | AG | AG | 1860 | |
| Db | 1801 | CAG | ATATGA | AGAG | AGTCTA | AGG | GAAG | AGCTG | CAAA | CTG | TA | AG | AG | 1860 | |

```
1861 CTGAGAAATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGCTTAATCAAGTATTTT 1920
1861 CTGAGAAATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGCTTAATCAAGTATTTT 1920
1921 GCAAACTGACATCAGATTTCTTAATGCTGTGCAAGACACATAATTCCTTAAATGAAC 1980
1921 GCAAACTGACATCAGATTTCTTAATGCTGTGCAAGACACATAATTCCTTAAATGAAC 1980
1981 ACTGCTATTTTAAATGAAACCTTTTCAATTCAGATTTTAAAGGGTAACCTTTT 2040
1981 ACTGCTATTTTAAATGAAACCTTTTCAATTCAGATTTTAAAGGGTAACCTTTT 2040
2041 ATTGCAATTTGCTGTGTTTCTATAAATGACATTTGTAATGCAATGCAAGCTTTTG 2100
2041 ATTGCAATTTGCTGTGTTTCTATAAATGACATTTGTAATGCAATGCAAGCTTTTG 2100
2101 AATGTGTAGTGTGCTGCTGTTCTGTGTCATAGTATCATCAAAAGTGGGTACAGTAAAGAG 2160
2101 AATGTGTAGTGTGCTGCTGTTCTGTGTCATAGTATCATCAAAAGTGGGTACAGTAAAGAG 2160
2161 CTTCAGCAATTTACTTTAACTCCCTCAACAGGTATACCTCAGTTCACAGGTTGTTAA 2220
2161 CTTCAGCAATTTACTTTAACTCCCTCAACAGGTATACCTCAGTTCACAGGTTGTTAA 2220
2221 TTATAAATTTGAAACACTAACAGAAATTTGAATAAATCAATGTCCTATTTTATAACAAGGT 2280
2221 TTATAAATTTGAAACACTAACAGAAATTTGAATAAATCAATGTCCTATTTTATAACAAGGT 2280
2281 TAATTCAAAATCTGTTGTTTATTAAGAAATGTTAAGTATGCTTTAGTGCCTAATAG 2340
2281 TAATTCAAAATCTGTTGTTTATTAAGAAATGTTAAGTATGCTTTAGTGCCTAATAG 2340
2341 TAAGTGGCTATTTTGAAGCAGTGTGTTTGTAGCTTTTCTTCTACTGCTTCTTAATTAGGGA 2400
2341 TAAGTGGCTATTTTGAAGCAGTGTGTTTGTAGCTTTTCTTCTACTGCTTCTTAATTAGGGA 2400
2401 AAACAAGTGTGCTGTTTGAATGGAAGAAATATGTTGTCACCTACCCCTACCCCTACTTAT 2460
2401 AAACAAGTGTGCTGTTTGAATGGAAGAAATATGTTGTCACCTACCCCTACCCCTACTTAT 2460
2461 ATCAAGTGTGCTGTTTGAATGGAAGAAATATGTTGTCACCTACCCCTACCCCTACTTAT 2520
2461 ATCAAGTGTGCTGTTTGAATGGAAGAAATATGTTGTCACCTACCCCTACCCCTACTTAT 2520
2521 CAGCCAGTAGAAGAAAGCTTAAACAGCTTTTCAACAGCTTTTCAACAGCTTTTCAACAGCT 2563
2521 CAGCCAGTAGAAGAAAGCTTAAACAGCTTTTCAACAGCTTTTCAACAGCTTTTCAACAGCT 2563
```

```
RESULT 2
AAC85299 standard; cDNA; 2563 BP.
AAC85299;
29-MAR-2001 (first entry)
Human activin receptor cDNA.
Mouse; human; rat; Xenopus; activin receptor; ligand binding domain;
transmembrane domain; receptor domain; serine kinase; TGF-beta;
transforming growth factor-beta; carcinogenesis; cancer; wound healing;
ss.
Homo sapiens.
Key Location/Qualifiers
CDS 71..1612
/*tag= a
/*product= "Activin receptor"
sig_peptide 71..127
/*tag= b
unsure 185..187
```

```
FT FT /*tag= C
FT FT /note= "Encodes Lys"
FT FT 344..346
FT FT /*tag= d
FT FT /note= "Encodes Val"
FT FT 932..934
FT FT /*tag= e
FT FT /note= "Encodes Gln"
XX XX
XX XX US6162896-A.
XX XX 19-DEC-2000.
XX XX 07-JUN-1995; 95US-00476123.
XX XX 10-MAY-1991; 91US-00698709.
XX XX 09-OCT-1991; 91US-00773229.
XX XX 08-MAY-1992; 92US-00880220.
XX XX 02-SEP-1994; 94US-00300584.
XX XX (SALK ) SALK INST BIOLOGICAL STUDIES.
XX XX Tsuchida K, Vale WW, Mathews LW;
XX XX WPI; 2001-090408/10.
XX XX P-PSDB; AAB47027.
XX XX Novel vertebrate activin receptor having extracellular ligand binding
XX XX domain, transmembrane domain and intracellular serine/threonine kinase
XX XX domain is useful for diagnosing and treating e.g. carcinogenesis, wound
XX XX healing.
XX XX Claim 1; Page; 33pp; English.
XX XX This sequence encodes a human-derived activin receptor. The activin
XX XX receptors of the invention comprise three distinct domains: an
XX XX extracellular, ligand binding domain, a hydrophobic, transmembrane domain
XX XX and an intracellular, receptor domain having serine kinase-like activity.
XX XX These proteins have binding affinity for at least one member of the
XX XX activin/TGF-beta superfamily of polypeptide growth factors. The activin
XX XX receptor proteins of the invention can be employed for a variety of
XX XX therapeutic uses, e.g. to block receptors. The presence of the soluble
XX XX proteins will compete with functional ligand for the receptor, preventing
XX XX the formation of a functional receptor-ligand complex, thereby blocking
XX XX the normal regulatory action of the complex. The receptor proteins are
XX XX useful for the diagnosis and therapeutic management of carcinogenesis,
XX XX wound healing, disorders of the immune, reproductive, or central nervous
XX XX systems. Note: This sequence is not shown in the specification but is
XX XX derived from the mouse-derived activin receptor coding sequence (see
XX XX AAC85297)
XX XX Sequence 2563 BP; 772 A; 484 C; 584 G; 720 T; 0 U; 3 Other;
XX XX Query Match 99.9%; Score 2561.2; DB 4; Length 2563;
XX XX Best Local Similarity 100.0%; Pred. No. 0;
XX XX Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGGAGAACCCAGGAACTGGATATCTAGCGAGAACTTCCTACGCTTCTCCGGCG 60
Db 1 CTCGAGGAGAACCCAGGAACTGGATATCTAGCGAGAACTTCCTACGCTTCTCCGGCG 60
Qy 61 CTCGGGAAATGGAGCTGCTGCAAAAGTGGCGTTCGGGCTCTTCTTATCTTCTGTC 120
Db 61 CTCGGGAAATGGAGCTGCTGCAAAAGTGGCGTTCGGGCTCTTCTTATCTTCTGTC 120
Qy 121 TTCAGTCTATACCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTG 180
Db 121 TTCAGTCTATACCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTG 180
Qy 181 GGAAGAAGACAGAACCAACACAGACTGGTGTGAACCTTGTATGGTGATAAAGATAACG 240
Db 181 GGAAGAAGACAGAACCAACACAGACTGGTGTGAACCTTGTATGGTGATAAAGATAACG 240
```

241 GGCACATTTGTTTGTCTACCTGGAAGAAATATTTCTGGTTCCATTGAAATAGTGAAGCAAGG 300
Db
241 GGCACATTTGTTTGTCTACCTGGAAGAAATATTTCTGGTTCCATTGAAATAGTGAAGCAAGG 300
Qy
301 TTGTTGGCTGGATGATATCAATCTGCTATGACAGACATGATTTGTGTNGAAGAAAAAGACAG 360
Db
301 TTGTTGGCTGGATGATCAATCTGCTATGACAGACATGATTTGTGTNGAAGAAAAAGACAG 360
Qy
361 CCCTGAAGTGTACTTTTGTGTGCTGAGGGCAATATGTAATGAAAGTTCTCTTATTT 420
Db
361 CCCTGAAGTGTACTTTTGTGTGCTGAGGGCAATATGTAATGAAAGTTCTCTTATTT 420
Qy
421 TCCGAGATGGGAAGTACACAGACCCCACTTCAAAATCCCTGTTACACCGAAGCCACCCCTATTA 480
Db
421 TCCGAGATGGGAAGTACACAGACCCCACTTCAAAATCCCTGTTACACCGAAGCCACCCCTATTA 480
Qy
481 CAACATTTCTGCTGTATTTCTTGTGTACCAATGTTAAATTTGACAGAAATGTCATTTGTC 540
Db
481 CAACATTTCTGCTGTATTTCTTGTGTACCAATGTTAAATTTGACAGAAATGTCATTTGTC 540
Qy
541 ATTTGGGTGTACAGACATACAGATGGCTACCCCTCTGTTACTTTGTTCTTACTCAAGA 600
Db
541 ATTTGGGTGTACAGACATACAGATGGCTACCCCTCTGTTACTTTGTTCTTACTCAAGA 600
Qy
601 CCCAGGACACCCCACTTCCCACTTACTAGGTTTGAAGCCATTGCAAGCTTGTAGAAAT 660
Db
601 CCCAGGACACCCCACTTCCCACTTACTAGGTTTGAAGCCATTGCAAGCTTGTAGAAAT 660
Qy
661 GAAAGCAAGGGGAAGATTTGGTTGTGTTCTGGAAGCCCACTGTTCAATGAATATGTGTC 720
Db
661 GAAAGCAAGGGGAAGATTTGGTTGTGTTCTGGAAGCCCACTGTTCAATGAATATGTGTC 720
Qy
721 TGTCAAAATATTTCCAATACAGACAAACAGTCTCTGGCAGAAATGAATATGAATCTATAG 780
Db
721 TGTCAAAATATTTCCAATACAGACAAACAGTCTCTGGCAGAAATGAATATGAATCTATAG 780
Qy
781 TCTACCTGGAATGAAGCATGAGACATACACTACAGTTCAATGGTGCAGAGAAAGGGCAC 840
Db
781 TCTACCTGGAATGAAGCATGAGACATACACTACAGTTCAATGGTGCAGAGAAAGGGCAC 840
Qy
841 CAGTGTGGATGTGGCTGTGGCTTAATACAGCATTTTCATGAAAAGGGCTCACTGTGAGA 900
Db
841 CAGTGTGGATGTGGCTGTGGCTTAATACAGCATTTTCATGAAAAGGGCTCACTGTGAGA 900
Qy
901 CTTTCTTAAGGCTAATGTGTCTCTTGAATPCARCTTTGTCAATTTGAGAAACCATGGC 960
Db
901 CTTTCTTAAGGCTAATGTGTCTCTTGAATPCARCTTTGTCAATTTGAGAAACCATGGC 960
Qy
961 TAGAGATTTGGCATATTTATGAGGATATACCTGGCTTAAAGATGGCCACCAAGCTGTC 1020
Db
961 TAGAGATTTGGCATATTTATGAGGATATACCTGGCTTAAAGATGGCCACCAAGCTGTC 1020
Qy
1021 AATCTCTCAGGAGACATCAAAAGTAAATGTGCTGTTGAAACAACTGACAGCTTG 1080
Db
1021 AATCTCTCAGGAGACATCAAAAGTAAATGTGCTGTTGAAACAACTGACAGCTTG 1080
Qy
1081 CATTGCTGACTTTGGGTGGCTTAAAGTTCGAGGCTGGCAAGTCTGCAAGTGAACCCCA 1140
Db
1081 CATTGCTGACTTTGGGTGGCTTAAAGTTCGAGGCTGGCAAGTCTGCAAGTGAACCCCA 1140
Qy
1141 TGGGCAAGTTGGTACCCGGAGGTATATGGCTCCAGAGGTGTTGGAGGGTGTATAACTT 1200
Db
1141 TGGGCAAGTTGGTACCCGGAGGTATATGGCTCCAGAGGTGTTGGAGGGTGTATAACTT 1200
Qy
1201 CAAAGGGAAGCATTTCTGAGGATAGATATGTAAGCCATGGATAGTCTTATGGGAAT 1260
Db
1201 CAAAGGGAAGCATTTCTGAGGATAGATATGTAAGCCATGGATAGTCTTATGGGAAT 1260
Qy
1261 GGCTTCTGCTGCTGCTGAGATGGACCGGTAGATGATGATACATGTTACCAATTTGAGGA 1320
Db
1261 GGCTTCTGCTGCTGCTGAGATGGACCGGTAGATGATGATACATGTTACCAATTTGAGGA 1320
Qy
1321 AGAAATTTGGCAGCATCCATCTCTTGAAGATATGAGGAAGTTGTTGTGCATAAAAAA 1380

1321 AGAAATTTGGCAGCATCCATCTCTTGAAGATATGAGGAAGTTGTTGTGCATAAAAAA 1380
Qy
1381 GAGGCTGTTTTAAGAGATTTATTGGCAGAAACATGAGGAATGCAATGCTCTGTGAAC 1440
Db
1381 GAGGCTGTTTTAAGAGATTTATTGGCAGAAACATGAGGAATGCAATGCTCTGTGAAC 1440
Qy
1441 GATAGAAGAAATGTTGGGATCATGATGAGAACCCAGTTTATCAGCTGGATGTGTAGTGA 1500
Db
1441 GATAGAAGAAATGTTGGGATCATGATGAGAACCCAGTTTATCAGCTGGATGTGTAGTGA 1500
Qy
1501 AAGAAATCTCAGATGCAAGACATAACAATATCATTAACAAGACATTTGTAACAGT 1560
Db
1501 AAGAAATCTCAGATGCAAGACATAACAATATCATTAACAAGACATTTGTAACAGT 1560
Qy
1561 GGTCAAAATGTTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGTGGCA 1620
Db
1561 GGTCAAAATGTTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGTGGCA 1620
Qy
1621 CGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTAAAGTAAGGAAAGTGC 1680
Db
1621 CGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTAAAGTAAGGAAAGTGC 1680
Qy
1681 TTAGTTGATTTCTGTGTGAATGAGTAGGATGCTCCAGACATGTAACGCAAGAGCC 1740
Db
1681 TTAGTTGATTTCTGTGTGAATGAGTAGGATGCTCCAGACATGTAACGCAAGAGCC 1740
Qy
1741 CTTGTGGAAGCATGGAATCTGGGAGATGGATCTGGGAACTTACTGCACTGCTGCAGCA 1800
Db
1741 CTTGTGGAAGCATGGAATCTGGGAGATGGATCTGGGAACTTACTGCACTGCTGCAGCA 1800
Qy
1801 CAGATATGAAGAGAGTCTAAGGGAAGCTGTAAGAACTGTAAGAACTTCTGAAATGTA 1860
Db
1801 CAGATATGAAGAGAGTCTAAGGGAAGCTGTAAGAACTGTAAGAACTTCTGAAATGTA 1860
Qy
1861 CTGGAAGAAATGTGGCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTT 1920
Db
1861 CTGGAAGAAATGTGGCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTT 1920
Qy
1921 GCAAACTGACATCAGATTTCTTAATGTCTGTGAGACACTAATTTCTTAAATGAACT 1980
Db
1921 GCAAACTGACATCAGATTTCTTAATGTCTGTGAGACACTAATTTCTTAAATGAACT 1980
Qy
1981 ACTGCTATTTTAAATGAAACTTTTCAATTCAGATTTTAAAGGGTAACTTTTT 2040
Db
1981 ACTGCTATTTTAAATGAAACTTTTCAATTCAGATTTTAAAGGGTAACTTTTT 2040
Qy
2041 ATTGCAATTTGCTGTTCTTATAAATGACTATTTGTAATGCCAATGACAGACTTTGTG 2100
Db
2041 ATTGCAATTTGCTGTTCTTATAAATGACTATTTGTAATGCCAATGACAGACTTTGTG 2100
Qy
2101 AATGTGTAGTGTGCTGCTGTTCTGTGTACATAGTCAATCAAGTGGGTACAGTAAAGAG 2160
Db
2101 AATGTGTAGTGTGCTGCTGTTCTGTGTACATAGTCAATCAAGTGGGTACAGTAAAGAG 2160
Qy
2161 CTTCCAAAGCATTTTAAACCTCCCTCAAGAGTATACCTCAGTTCCAGGTTGTAA 2220
Db
2161 CTTCCAAAGCATTTTAAACCTCCCTCAAGAGTATACCTCAGTTCCAGGTTGTAA 2220
Qy
2221 TTATAAATTTGAACACTTAACAGATTTGTAATTAATCAGTCCATGTTTATAACAAGT 2280
Db
2221 TTATAAATTTGAACACTTAACAGATTTGTAATTAATCAGTCCATGTTTATAACAAGT 2280
Qy
2281 TAATTAACAATTTCACTGTTTATTTAAGAAAAATGGTAAAGCTATGTTAGTCCCAATAG 2340
Db
2281 TAATTAACAATTTCACTGTTTATTTAAGAAAAATGGTAAAGCTATGTTAGTCCCAATAG 2340
Qy
2341 TAAGTGGCTATTTGTAAGAGCTGTTTGTAGCTTTTCTTACTGCTGCTGTAATTTAGGA 2400
Db
2341 TAAGTGGCTATTTGTAAGAGCTGTTTGTAGCTTTTCTTACTGCTGCTGTAATTTAGGA 2400
Qy
2401 AAACAAGTGTGTTTGAATGGAAGAAATATGTTGTGTCACCTACCCCTTACTTAT 2460

Db 2401 AACCAAGTGTCTTTGAAATGGAAAGAAATATGTGTACCCCTACCCCTACTTAT 2460
2y 2461 ATCAAGGTCACAAATATTTCTTTCCATTTCATTAAGACAGACACTTTGAAACCCCTAAATTA 2520
Db 2461 ATCAAGGTCACAAATATTTCTTTCCATTTCATTAAGACAGACACTTTGAAACCCCTAAATTA 2520
2y 2521 CAAGCCAGTAGAAGAAAGCTAAACACACGCTTTAATAATAGCC 2563
Db 2521 CAAGCCAGTAGAAGAAAGCTAAACACACGCTTTAATAATAGCC 2563

RESULT 3

AAK98725

ID AAK98725 standard; cDNA; 2563 BP.

XX AC AAK98725;

XX 07-MAY-2002 (first entry)

XX cDNA encoding a human-derived activin receptor.

XX Activin receptor; cloning; recombinant; TGF-beta; ligand-binding;
KW superfamily; trans-membrane; receptor domain; serine kinase; diagnosis;
KW therapeutic management; carcinogenesis; wound healing; protein therapy;
KW immune; reproductive; central nervous system; activin-dependent tumour;
KW brain neuron; abortion; livestock; twinning; probe; agonist; cytostatic;
KW wound healing; transplant organ rejection; vulnerable; immunosuppressive;
KW transforming growth factor-beta; human; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

5'UTR 1..70

FT /*tag= a

FT 71..1612

FT /*tag= b

FT /product= "Human-derived activin receptor protein"

FT 3'UTR 1613..2563

FT /*tag= c

XX US2001039036-A1.

XX 08-NOV-2001.

XX 19-DEC-2000; 2000US-00742684.

XX 10-MAY-1991; 91US-00698709.

XX 09-OCT-1991; 91US-00773229.

XX 08-MAY-1992; 92US-00880220.

XX 02-SEP-1994; 94US-00300584.

XX 07-JUN-1995; 95US-00476123.

XX (SALK) SALK INST BIOLOGICAL STUDIES.

XX Mathews LS, Vale WW, Tsuchida K;

XX WPI; 2002-040721/05.

XX P-PSDB; AAC14127.

XX New receptor proteins having an extracellular ligand-binding domain, a
PT hydrophobic trans-membrane domain, and an intracellular receptor domain,
PT useful for diagnosing or treating carcinogenesis, wound healing or immune
PT disorders.

XX Example 11; Page; 33pp; English.

XX This polynucleotide sequence represents the cDNA encoding a human-derived
CC activin receptor. The invention relates to cloning and recombinant
CC production of receptor(s) of the activin/TGF-beta (transforming growth
CC factor-beta) superfamily. The invention has identified and characterised
CC members of a new superfamily of receptor proteins which comprise three
CC distinct domains: an extracellular, ligand-binding domain, a hydrophobic,
CC trans-membrane domain, and an intracellular, receptor domain having

CC serine kinase-like activity. The receptor proteins and antibodies to
CC these proteins are useful in the diagnosis and therapeutic management of
CC carcinogenesis, wound healing, disorders of the immune, reproductive, or
CC central nervous systems. The receptor proteins of the invention can be
CC used in protein therapy. These may further be used to diagnose or treat
CC activin-dependent tumours, enhance the survival of brain neurons, induce
CC abortion in livestock and other domesticated animals, and induce twinning
CC in livestock and other domesticated animals. The DNAs are useful as
CC probes for identifying additional members of the superfamily of receptor
CC proteins, and as coding sequences which can be used for the recombinant
CC expression of the receptor proteins. Agonists for TGF-beta specific
CC receptors can be used to stimulate wound healing, to suppress growth of
CC TGF-beta sensitive tumours, or to suppress immune response (thus prevent
CC rejection of transplant organ). The receptor proteins of the invention
CC have cytostatic, vulnerary, and immunosuppressive activity. NOTE: This
CC sequence is not shown in the specification. It has been created from the
CC sequence AAK98720 and information provided on page 12 of the
CC specification

XX SQ Sequence 2563 BP; 772 A; 484 C; 584 G; 720 T; 0 U; 3 Other;

Query Match 99.9%; Score 2561.2; DB 6; Length 2563;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGGAAGACCCAGGGAACCTGGATATCTAGCAGAACTTCTAGCGCTTCTCCGGCG 60

Db 1 CTCGAGGAAGACCCAGGGAACCTGGATATCTAGCAGAACTTCTAGCGCTTCTCCGGCG 60

Qy 61 CCTCGGAAAAATGGGAGCTGCTGCAAAAGTTGGGGTTCGCCGTCTTTCTTATCTCTTGCTC 120

Db 61 CCTCGGAAAAATGGGAGCTGCTGCAAAAGTTGGGGTTCGCCGTCTTTCTTATCTCTTGCTC 120

Qy 121 TTCAGGTGCTATACCTTGGCAGATCAGAACTCAGGAGTGCTTTTCTTTAATGCTAATG 180

Db 121 TTCAGGTGCTATACCTTGGCAGATCAGAACTCAGGAGTGCTTTTCTTTAATGCTAATG 180

Qy 181 GGAAGAAGACAGAACCAACGAGCTGGTGTGAACCTTGTATGTTGATGAAGATAACG 240

Db 181 GGAAGAAGACAGAACCAACGAGCTGGTGTGAACCTTGTATGTTGATGAAGATAACG 240

Qy 241 GGCACATTTGTTTGTCTGCTGGAAGAAATATTTCTGGTTCCATTGAAATAGTGAACAAG 300

Db 241 GGCACATTTGTTTGTCTGCTGGAAGAAATATTTCTGGTTCCATTGAAATAGTGAACAAG 300

Qy 301 TTGTTGGCTGGATGATATCAACTGCTATGACAGGACTGATTTGTGNGAAAAAAGACAG 360

Db 301 TTGTTGGCTGGATGATATCAACTGCTATGACAGGACTGATTTGTGNGAAAAAAGACAG 360

Qy 361 CCTGAAAGTGTAATTTTGTGCTGAGGGAATATGTAATGAAAGTTCTCTTATTT 420

Db 361 CCTGAAAGTGTAATTTTGTGCTGAGGGAATATGTAATGAAAGTTCTCTTATTT 420

Qy 421 TCCGGAGATGGAAGTCACACAGCCCACTTCAATCTCTTACACCGAGCCACCTATTA 480

Db 421 TCCGGAGATGGAAGTCACACAGCCCACTTCAATCTCTTACACCGAGCCACCTATTA 480

Qy 481 CAACATTTCTGTGTAATTTCTTGTGTACCACTAAATGTTAATGTCAGGAATTTGCTGTC 540

Db 481 CAACATTTCTGTGTAATTTCTTGTGTACCACTAAATGTTAATGTCAGGAATTTGCTGTC 540

Qy 541 ATTTTGGGTGTACAGACATCAAGATGGCTACCTTCCCTGTACTTGTTCCTACTCAAGA 600

Db 541 ATTTTGGGTGTACAGACATCAAGATGGCTACCTTCCCTGTACTTGTTCCTACTCAAGA 600

Qy 601 CCCAGACACACCCCACTTCCCTCATTTAGGGTTGAAGCCATTGACGCTGTAGAGT 660

Db 601 CCCAGACACACCCCACTTCCCTCATTTAGGGTTGAAGCCATTGACGCTGTAGAGT 660

Qy 661 GAAAGCAAGGGAGAGATTGGTGTCTGGAAGCCAGTTGCTCAATGATATGTCG 720

Db 661 GAAAGCAAGGGAGAGATTGGTGTCTGGAAGCCAGTTGCTCAATGATATGTCG 720

| | | | |
|----|------|--|------|
| Qy | 721 | TGTCAAAATATTTTCCAAATACAGACAAACAGTCCTGCGAGAATGAATATGAAGTCTATAG | 780 |
| Db | 721 | TGTCAAAATATTTTCCAAATACAGACAAACAGTCCTGCGAGAATGAATATGAAGTCTATAG | 780 |
| Qy | 781 | TCTACTCGGAATGAAGCATGAGAAACATACCTACAGTTCACTTGGTCGAGAGAAAGAGGCAC | 840 |
| Db | 781 | TCTACTCGGAATGAAGCATGAGAAACATACCTACAGTTCACTTGGTCGAGAGAAAGAGGCAC | 840 |
| Qy | 841 | CAGTCGTGATGGGACCTGTGGCTAAACACAGCAATTCATCAAAAGGACTCACTCTCAGA | 900 |
| Db | 841 | CAGTCGTGATGGGACCTGTGGCTAAACACAGCAATTCATCAAAAGGACTCACTCTCAGA | 900 |
| Qy | 901 | CTTTCTTAAGGCTAAATGTGGTCTCTTGGAAATCARCTTTGTCTATATTGCAGAAACCATGGC | 960 |
| Db | 901 | CTTTCTTAAGGCTAAATGTGGTCTCTTGGAAATCARCTTTGTCTATATTGCAGAAACCATGGC | 960 |
| Qy | 961 | TAGAGGATGGCATATTTATCATGAGGATATACCTGGCTTAAAGATGCCCACAAGCCTGC | 1020 |
| Db | 961 | TAGAGGATGGCATATTTATCATGAGGATATACCTGGCTTAAAGATGCCCACAAGCCTGC | 1020 |
| Qy | 1021 | AATCTCTCACAGGACATCAAAAGTAAAAATGTGCTGTTGAAAAACAATCTGCACAGCTTG | 1080 |
| Db | 1021 | AATCTCTCACAGGACATCAAAAGTAAAAATGTGCTGTTGAAAAACAATCTGCACAGCTTG | 1080 |
| Qy | 1081 | CATTGCTGACTTTGGTTGGCCTTAAAGTTGAGGCTGGCAAGTCTGCAGGTGACACCCA | 1140 |
| Db | 1081 | CATTGCTGACTTTGGTTGGCCTTAAAGTTGAGGCTGGCAAGTCTGCAGGTGACACCCA | 1140 |
| Qy | 1141 | TGGCAGGTTGTTACCCGAGGTATATGGTCCAGAGGTGTGGAGGTTGCTATAAACTT | 1200 |
| Db | 1141 | TGGCAGGTTGTTACCCGAGGTATATGGTCCAGAGGTGTGGAGGTTGCTATAAACTT | 1200 |
| Qy | 1201 | CCAAAGGACGCAATTTCTGAGGATAGATATGTACGCCATGGGATAGTCTCTATGGGAAT | 1260 |
| Db | 1201 | CCAAAGGACGCAATTTCTGAGGATAGATATGTACGCCATGGGATAGTCTCTATGGGAAT | 1260 |
| Qy | 1261 | GGCTTCTCGTGTGCACTGTGCAGATGGAACCCGTAGATGAGTACATGTTACATTTGAGGA | 1320 |
| Db | 1261 | GGCTTCTCGTGTGCACTGTGCAGATGGAACCCGTAGATGAGTACATGTTACATTTGAGGA | 1320 |
| Qy | 1321 | AGAAATGGCCAGCATCCATCTCTTGAAGATATGCGAGGAAGTTGTTGTGCATAAAAAAAA | 1380 |
| Db | 1321 | AGAAATGGCCAGCATCCATCTCTTGAAGATATGCGAGGAAGTTGTTGTGCATAAAAAAAA | 1380 |
| Qy | 1381 | GAGGCTGTGTTTAAAGAGATTATGGCGAGAAACATGCAGGAATGGCAATGCTCTGTGAAC | 1440 |
| Db | 1381 | GAGGCTGTGTTTAAAGAGATTATGGCGAGAAACATGCAGGAATGGCAATGCTCTGTGAAC | 1440 |
| Qy | 1441 | GATAGAAGAAATGTTGGGATCATGATGACAGAAGCCAGGTTATCAGCTGATGTGAGTGA | 1500 |
| Db | 1441 | GATAGAAGAAATGTTGGGATCATGATGACAGAAGCCAGGTTATCAGCTGATGTGAGTGA | 1500 |
| Qy | 1501 | AAGAAATTAATCAGATCAAAAGACTAAACAAATATCATTACTACAGAGGACATTTGAACAGT | 1560 |
| Db | 1501 | AAGAAATTAATCAGATCAAAAGACTAAACAAATATCATTACTACAGAGGACATTTGAACAGT | 1560 |
| Qy | 1561 | GGTCACAAATGGTGACAAAATGTTGACTTTCCTCCAAAGAAATCTAGTCTATGATGGTGGCA | 1620 |
| Db | 1561 | GGTCACAAATGGTGACAAAATGTTGACTTTCCTCCAAAGAAATCTAGTCTATGATGGTGGCA | 1620 |
| Qy | 1621 | CCGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTTAAGCTTAAGGAAAGTGC | 1680 |
| Db | 1621 | CCGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTTAAGCTTAAGGAAAGTGC | 1680 |
| Qy | 1681 | TTAGTTGATTTTCTGTGTGAAATCAGTAGGATGCTCCAGACATGTTACGCAAGCAGCCC | 1740 |
| Db | 1681 | TTAGTTGATTTTCTGTGTGAAATCAGTAGGATGCTCCAGACATGTTACGCAAGCAGCCC | 1740 |
| Qy | 1741 | CTTTGTGGAAGCATGGAATCTGGGAGATGGATCTCTGGGAAACTTACTGCACTCGTCTGCAGCA | 1800 |
| Db | 1741 | CTTTGTGGAAGCATGGAATCTGGGAGATGGATCTCTGGGAAACTTACTGCACTCGTCTGCAGCA | 1800 |
| Qy | 1801 | CAGATATGAAGAGGAGTCTTAAGGGAAAAGCTGCACAACTGTAAAGAACTTCTGAAAATGTA | 1860 |

| | | | |
|----------|-----------------------------------|---|------|
| DB | 1801 | CAGATATGAAGAGGAGCTCTAAGGGGAAAAGCTGCAAACTGTAAAGAACTTCTGAAAAATGTA | 1861 |
| QY | 1861 | CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAAATCAAGTATTTT | 1920 |
| DB | 1861 | CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAAATCAAGTATTTT | 1920 |
| QY | 1921 | GCAAACTCACATCAGATTTCTTAATCTGTCTGTCAGAGACACTAAATCTCTTAAATGAACT | 1980 |
| DB | 1921 | GCAAACTGACATCAGATTTCTTAATCTGTCTGTCAGAGACACTAAATCTCTTAAATGAACT | 1980 |
| QY | 1981 | ACTGCTATTTTTTTTTTAAATGAAAAACCTTTTCATTTTCAGATTTTAAAAAGGGTAACTTTTT | 2040 |
| DB | 1981 | ACTGCTATTTTTTTTTTAAATGAAAAACCTTTTCATTTTCAGATTTTAAAAAGGGTAACTTTTT | 2040 |
| QY | 2041 | ATTGCATTTGCTGTTGTTTCTATAAATGACATTAATGTAATGCAACATGACACAGCTTTGTC | 2100 |
| DB | 2041 | ATTGCATTTGCTGTTGTTTCTATAAATGACATTAATGTAATGCAACATGACACAGCTTTGTC | 2100 |
| QY | 2101 | AATGTTGATGTGCTGCTGTTCTGTGTACATAGTCATCAAAAGTGGGTACAGTAAAGAGG | 2160 |
| DB | 2101 | AATGTTGATGTGCTGCTGTTCTGTGTACATAGTCATCAAAAGTGGGTACAGTAAAGAGG | 2160 |
| QY | 2161 | CTTCCAAGCATTTACTTTTAACTCCCTCAACAGGTATACCTCAGTTCACGGTTGTATAA | 2220 |
| DB | 2161 | CTTCCAAGCATTTACTTTTAACTCCCTCAACAGGTATACCTCAGTTCACGGTTGTATAA | 2220 |
| QY | 2221 | TTATAAAATGAAAAACACTAACAGAAATTTGAATAAATCAGTCCATGTTTTTATAACAAAGT | 2280 |
| DB | 2221 | TTATAAAATGAAAAACACTAACAGAAATTTGAATAAATCAGTCCATGTTTTTATAACAAAGT | 2280 |
| QY | 2281 | TAATTACAAATTCACGTGTGTTATTTAAGAAAAAATGGTAAGCTATGCTTAGTGCCCAATAG | 2340 |
| DB | 2281 | TAATTACAAATTCACGTGTGTTATTTAAGAAAAAATGGTAAGCTATGCTTAGTGCCCAATAG | 2340 |
| QY | 2341 | TAAGTGGCTATTGTTAAGCAGTGTGTTTTAGCTTTTCTCTACTGCTGTCTAATTTAGGGA | 2400 |
| DB | 2341 | TAAGTGGCTATTGTTAAGCAGTGTGTTTTAGCTTTTCTCTACTGCTGTCTAATTTAGGGA | 2400 |
| QY | 2401 | AAACAAGTCTGTCTTTGAAATGGAAAGAAATATGGTGTCAACCTACCCCTACCTACTTAT | 2460 |
| DB | 2401 | AAACAAGTCTGTCTTTGAAATGGAAAGAAATATGGTGTCAACCTACCCCTACCTACTTAT | 2460 |
| QY | 2461 | ATCAAGGTCCTCAAAATATTTCTTTTCCATTTCAAAGACAGCACTTTGAAAAACCTTAAATTA | 2520 |
| DB | 2461 | ATCAAGGTCCTCAAAATATTTCTTTTCCATTTCAAAGACAGCACTTTGAAAAACCTTAAATTA | 2520 |
| QY | 2521 | CAAGCCAGTAGAGAAAAGCTAAAACACGCTTTTACAAATAGCC | 2563 |
| DB | 2521 | CAAGCCAGTAGAGAAAAGCTAAAACACGCTTTTACAAATAGCC | 2563 |
| RESULT 4 | | | |
| AAQ31912 | | | |
| ID | AAQ31912 standard; cDNA; 2563 BP. | | |
| XX | AC | AAQ31912; | |
| XX | DT | 25-MAR-2003 (revised) | |
| XX | DT | 19-APR-1993 (first entry) | |
| XX | DE | Human activin receptor. | |
| XX | KW | Activin receptor; mouse; Xenopus; human; extracellular; ligand binding; | |
| XX | KW | hydrophobic; trans-membrane; intracellular; receptor; domain; | |
| XX | KW | serine kinase-like; activity; probe; superfamily; secretion signal; | |
| XX | KW | golgi membrane; diagnosis; treatment; activin-dependent tumour; brain; | |
| XX | KW | neuron; abortion; twinning; wound healing; TGF-beta; immune response; | |
| XX | KW | liver regeneration; ss. | |
| OS | Homo sapiens. | | |
| FH | Key | Location/Qualifiers | |

CDS 71..1612
/*tag= a
WO9220793-A1.
26-NOV-1992.
08-MAY-1992; 92WO-US003825.
10-MAY-1991; 91US-00698709.
09-OCT-1991; 91US-00773229.
(SALK) SALK INST BIOLOGICAL STUDIES.
Mathews LS, Vale WW;
WPI; 1992-415771/50.
P-PSDB; AAR29583.
New member of activin-transforming growth factor beta super-family - for
diagnosis and treatment of cancer and disorders of the immune,
reproductive or central nervous system.
Disclosure; Page 40; 68pp; English.
The sequences given in AA031910-12 encode activin receptors from mouse,
Xenopus and human respectively. Each of the proteins encoded by these
sequences comprise three distinct domains; an extracellular, ligand
binding domain, a hydrophobic, trans-membrane domain and an
intracellular, receptor domain having serine kinase-like activity. These
DNA sequences can be used as probes for the identification of additional
members of this superfamily of receptor molecules. The proteins may
further comprise a second hydrophobic domain at the amino terminal which
comprises a secretion signal sequence which promotes the intracellular
transport of the initially expressed receptor protein across the golgi
membrane. The receptor proteins can be used to develop agents for the
diagnosis and/or treatment of eg. activin-dependent tumours, for
enhancing the survival of brain neurons, for inducing abortion or
twinning in livestock, for stimulating wound healing, for suppression of
growth of TGF-beta sensitive tumours, for suppressing immune response,
for promoting liver regeneration and for stimulating some immune
responses. (Updated on 25-MAR-2003 to correct PN field.)
Sequence 2563' BP; 773 A; 486 C; 582 G; 719 T; 0 U; 3 Other;
Query Match 99.7%; Score 2556.4; DB 2; Length 2563;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2560; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 CTCGGAGAGACCCAGGAACTGGATATCTAGCAGAACTTCTACGGCTTCTCGGGG 60
1 CTCGGAGAGACCCAGGAACTGGATATCTAGCAGAACTTCTACGGCTTCTCGGGG 60
61 CCTCGGAAATGGAGCTGTCGAAAGTTGGCGTTGCGCGTTCTTCTATCTCTGCTC 120
61 CCTCGGAAATGGAGCTGTCGAAAGTTGGCGTTGCGCGTTCTTCTATCTCTGCTC 120
121 TTCAGGTGCTATCTTGGCAGATCAGAACTCAGAGTGTCTTTCTTAACTCTAATTG 180
121 TTCAGGTGCTATCTTGGCAGATCAGAACTCAGAGTGTCTTTCTTAACTCTAATTG 180
181 GGAAAGACAGAACCAACACAGACTGGTGTGAACTTCTATGGTGTGATAAAGATAACG 240
181 GGAAAGACAGAACCAACACAGACTGGTGTGAACTTCTATGGTGTGATAAAGATAACG 240
241 CGGACATGTTTGGTACTCTGGAAGATATTTCTGTTCCATGAAATAGTGAAGCAAG 300
241 CGGACATGTTTGGTACTCTGGAAGATATTTCTGTTCCATGAAATAGTGAAGCAAG 300
301 TTGTTGCTGATGATATCACTGCTATGACAGGACTGATGTTGTTGAAAAAAGACAG 360
301 TTGTTGCTGATGATATCACTGCTATGACAGGACTGATGTTGTTGAAAAAAGACAG 360

361 CCTGAAAGTGTACTTTTGTGCTGTGAGGGCAATATGTGTAAAGAGTCTCTTATTT 420
361 CCTGAAAGTGTACTTTTGTGCTGTGAGGGCAATATGTGTAAAGAGTCTCTTATTT 420
421 TCCGAGATGGAAGTCAACAGCCACCTTCAATCCTGTTACACCGAAGCCACCTATT 480
421 TCCGAGATGGAAGTCAACAGCCACCTTCAATCCTGTTACACCGAAGCCACCTATT 480
481 CAACATTCTGCTGATTCTTTGGTACCACTAATGTTAAATGCGAGGAATTGTCTGTC 540
481 CAACATTCTGCTGATTCTTTGGTACCACTAATGTTAAATGCGAGGAATTGTCTGTC 540
541 ATTTTGGGTGACAGACATCAAGATGCGCTACCTCTCTGTACTTGTTCCTACTCA 600
541 ATTTTGGGTGACAGACATCAAGATGCGCTACCTCTCTGTACTTGTTCCTACTCA 600
601 CCCAGGACCAACCCACCTTCCCCTTACTAGGTTGAAGCCATTCAGCTGTTAGAA 660
601 CCCAGGACCAACCCACCTTCCCCTTACTAGGTTGAAGCCATTCAGCTGTTAGAA 660
661 GAAAGCAAGGGAAGATTTGGTGTGCTGGAAGCCCAAGTTCCTCAATGAATATG 720
661 GAAAGCAAGGGAAGATTTGGTGTGCTGGAAGCCCAAGTTCCTCAATGAATATG 720
721 TGTCAAAATATTTCCAAATACAGGACAAACAGTCCTGGCAGAAATGAATATGA 780
721 TGTCAAAATATTTCCAAATACAGGACAAACAGTCCTGGCAGAAATGAATATGA 780
781 TCTACCTGGAATGAAGCATGAAACATACACTACAGTTCATTGTCGAGAGAAAG 840
781 TCTACCTGGAATGAAGCATGAAACATACACTACAGTTCATTGTCGAGAGAAAG 840
841 CAGTGTGATGTGACCTGTGCTTAATCACAGCATTTCAATGAAAGGGCTCACTGT 900
841 CAGTGTGATGTGACCTGTGCTTAATCACAGCATTTCAATGAAAGGGCTCACTGT 900
901 CTTTCTTAAAGCTTAATGTTGCTCTTGGAAATCACTTTGTCATATTGCAAAAC 960
901 CTTTCTTAAAGCTTAATGTTGCTCTTGGAAATCACTTTGTCATATTGCAAAAC 960
961 TAGAGGATGTCATATTTACATGAGATATACCTGCTTAAAGAGTGGCCACAG 1020
961 TAGAGGATGTCATATTTACATGAGATATACCTGCTTAAAGAGTGGCCACAG 1020
1021 AATCTCTCAGAGGACATCAAAAGTAAAGTGTGTTGAAAAACAATCTGACAG 1080
1021 AATCTCTCAGAGGACATCAAAAGTAAAGTGTGTTGAAAAACAATCTGACAG 1080
1081 CATTGCTGACTTTGGGTTGGCCTTAAAGTTCGAGGCTGCAAGTCTGCAAGTCA 1140
1081 CATTGCTGACTTTGGGTTGGCCTTAAAGTTCGAGGCTGCAAGTCTGCAAGTCA 1140
1141 TGGCAGGTTGGTACCCGAGGTATATGGCTCCAGAGTGTGGAGGGTCTATAA 1200
1141 TGGCAGGTTGGTACCCGAGGTATATGGCTCCAGAGTGTGGAGGGTCTATAA 1200
1201 CCAGAGGACGCAATTTCTGAGGATAGATATGACCCATGGGATTAAGTCTTA 1260
1201 CCAGAGGACGCAATTTCTGAGGATAGATATGACCCATGGGATTAAGTCTTA 1260
1261 GGCTTCTGTTGCTGCTGAGATGAGCCCGTAGATGAGTACATGTTACCATTT 1320
1261 GGCTTCTGTTGCTGCTGAGATGAGCCCGTAGATGAGTACATGTTACCATTT 1320
1321 AGAATTTGGCCAGCATCCATCTCTTGAAGATATGCAAGGAGTTGTTGTGAT 1380
1321 AGAATTTGGCCAGCATCCATCTCTTGAAGATATGCAAGGAGTTGTTGTGAT 1380
1381 GAGGCTCTTTTAAAGATATTATGCGAGAAACATGCGAGGAATGCGCAATGCT 1440
1381 GAGGCTCTTTTAAAGATATTATGCGAGAAACATGCGAGGAATGCGCAATGCT 1440
1441 GATAGAAAGATGTTGGGATCATGATGACAGAGCCAGGTTATCAGCTGGATGT 1500

Db 1441 GATAGAAGAAATGTTGGGATCATGATGAGAGGAGGTTATCAGCTGGATGTGAGTGA 1500
Qy 1501 AAGAAATTAATCAAGTCAAGAAAGTAAACAAATATCACTTACTACAGAGGACATTTGTAACAGT 1560
Db 1501 AAGAAATTAATCAAGTCAAGAAAGTAAACAAATATCACTTACTACAGAGGACATTTGTAACAGT 1560
Qy 1561 GGTCAAAATGTTGACAAATGTTGACATTTCTCCCAAGAAATCTAGTCTATGATGTTGCA 1620
Db 1561 GGTCAAAATGTTGACAAATGTTGACATTTCTCCCAAGAAATCTAGTCTATGATGTTGCA 1620
Qy 1621 CGTCTGTACACATGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAGTGC 1680
Db 1621 CGTCTGTACACATGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAGTGC 1680
Qy 1681 TTAGTTGATTTCTGTGTAATGAGTGAATGCTCCAGGACATGTACGCAAGCAGGCC 1740
Db 1681 TTAGTTGATTTCTGTGTAATGAGTGAATGCTCCAGGACATGTACGCAAGCAGGCC 1740
Qy 1741 CTTGTGGAAGCATGATCTGGAGATGATCTCTGGAACTTACTGCACTGCTGCAGCA 1800
Db 1741 CTTGTGGAAGCATGATCTGGAGATGATCTCTGGAACTTACTGCACTGCTGCAGCA 1800
Qy 1801 CAGATATGAGGAGGCTTAAGGGAAGAGTCTGAACTGCAAGACATCTTCTGAAATGTA 1860
Db 1801 CAGATATGAGGAGGCTTAAGGGAAGAGTCTGAACTGCAAGACATCTTCTGAAATGTA 1860
Qy 1861 CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTGGACCTGGCTAACTCAAGTATT 1920
Db 1861 CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTGGACCTGGCTAACTCAAGTATT 1920
Qy 1921 GCAAAATGACATCAGATTTCTTAATGCTGTCTGAGAGACACTAATCTCTTAAATGAACT 1980
Db 1921 GCAAAATGACATCAGATTTCTTAATGCTGTCTGAGAGACACTAATCTCTTAAATGAACT 1980
Qy 1981 ACTGCTATTTTCTTAATGAAACAACTTTTCAATTCAGATTTTAAAGGGTAACTTTT 2040
Db 1981 ACTGCTATTTTCTTAATGAAACAACTTTTCAATTCAGATTTTAAAGGGTAACTTTT 2040
Qy 2041 ATTGCATTTGCTGTTTCTTAATGATGATTAATGATTAATGATTAATGATTAATGATTAATG 2100
Db 2041 ATTGCATTTGCTGTTTCTTAATGATGATTAATGATTAATGATTAATGATTAATGATTAATG 2100
Qy 2101 AATGCTAGTGTGCTGTTCTGTGTACATGATGATGATGATGATGATGATGATGATGATGATG 2160
Db 2101 AATGCTAGTGTGCTGTTCTGTGTACATGATGATGATGATGATGATGATGATGATGATGATG 2160
Qy 2161 CTTCCAGCATTAATCTTAACCTCCCTCAACAGGATATACCTCAGTTCCACGCTTGTAAA 2220
Db 2161 CTTCCAGCATTAATCTTAACCTCCCTCAACAGGATATACCTCAGTTCCACGCTTGTAAA 2220
Qy 2221 TTATAAAATGAAACACTAACAAGATTTGAATTAATCAGTCCATGTTTATTAACAAGGT 2280
Db 2221 TTATAAAATGAAACACTAACAAGATTTGAATTAATCAGTCCATGTTTATTAACAAGGT 2280
Qy 2281 TAATTACAAATCTGCTGTTTATTAAGAAATGTAAGTAACTGCTAGTCCCAATAG 2340
Db 2281 TAATTACAAATCTGCTGTTTATTAAGAAATGTAAGTAACTGCTAGTCCCAATAG 2340
Qy 2341 TAAGTGGCTATTTGTAAGCAGTGTGTTAGCTTTTCTTCTACTGCTGCTGTTTATTAAGGCA 2400
Db 2341 TAAGTGGCTATTTGTAAGCAGTGTGTTAGCTTTTCTTCTACTGCTGCTGTTTATTAAGGCA 2400
Qy 2401 AAACAAGTGTGCTTTTGAATGGAAGAAATATGTTGTCACCTTACCCCTACCCCTATCTTAT 2460
Db 2401 AAACAAGTGTGCTTTTGAATGGAAGAAATATGTTGTCACCTTACCCCTACCCCTATCTTAT 2460
Qy 2461 ATCAAGTCCCAAAATATCTTTTCCATTTCAAAGACAGCATTTGAAACCCCTAAATTA 2520
Db 2461 ATCAAGTCCCAAAATATCTTTTCCATTTCAAAGACAGCATTTGAAACCCCTAAATTA 2520
Qy 2521 CAGCCAGTAGAAGAAAGCTTAAACACGCTTTTCAATAGCC 2563

Db 2521 CAGCCAGTAGAAGAAAGCTTAAACACGCTTTTCAATAGCC 2563
RESULT 5
AA222683
ID AAX22683 standard; cDNA; 2563 BP.
XX AAX22683;
AC AAX22683;
DT 27-MAY-1999 (first entry)
XX Mouse activin receptor pre-cursor cDNA.
DE Activin receptor; activin/TGF-superfamily; receptor protein; function;
KW activity; modulate; treatment; carcinogenesis; wound healing; fertility;
KW immune system disorder; central nervous system disorder; mouse;
KW reproductive system; ss.
XX Mus sp.
OS
XX Key Location/Qualifiers
FH 71..1612
CDS /*tag= a
FT /product= "activin receptor precursor"
FT
XX US885794-A.
XX 23-MAR-1999. 94US-00300584.
XX 02-SEP-1994; 91US-00698709.
XX 10-MAY-1991; 91US-00773229.
XX 09-OCT-1991; 92US-00880220.
XX 08-MAY-1992; 92US-00880220.
XX (SALK) SALK INST BIOLOGICAL STUDIES.
PA Vale WW, Mathews LS;
XX WPI; 1999-228534/19.
XX P-PSDB; AAW3203.
XX Isolated nucleic acid molecules encoding vertebrate activin receptor polypeptides - useful as probes for detecting similar sequences and for investigating the function of the receptor in conditions such as carcinogenesis, wound healing and disorders of the immune, central nervous and reproductive systems.
XX Claim 1a; Col 21-26; 28pp; English.
XX This sequence encodes a mouse activin receptor polypeptide. The nucleic acid molecules of the invention are useful as probes for the identification of additional members of the activin/TGF-superfamily of receptor proteins, and the coding sequences can be used for the recombinant expression of the receptor proteins or functional fragments of them. They may also be used to study the function and activity of activin receptor polypeptides in cells and to identify agents which will modulate activin receptor expression and activity for use in treating conditions such as carcinogenesis, wound healing, disorders of the immune or central nervous systems and especially the reproductive system (where they may be used to control fertility in humans, domestic and commercial animals)

Query Match 99.7%; Score 2556.4; DB 2; Length 2563;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2557; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 1 CTCGAGAGAACCCAGGAACTGGATATCTAGCAGAACTTCTACGGCTTCTCGGCG 60
Db 1 CTCGAGAGAACCCAGGAACTGGATATCTAGCAGAACTTCTACGGCTTCTCGGCG 60

| | | |
|------|--|------|
| 61 | CTCTGGGAAAATGGGAGCTGCTGCAAAAGTTGGCGTTTCGCCGTCCTTTCTTATCTCTTGCTC | 120 |
| iy | | |
| 61 | CTCTGGGAAAATGGGAGCTGCTGCAAAAGTTGGCGTTTCGCCGTCCTTTCTTATCTCTTGCTC | 120 |
| ib | | |
| 121 | TTCAAGTGTCTATCTCTTGGCAGATCAGAACTCAGGAGTGTCTTTTCTTTTAATGCTTAATG | 180 |
| iy | | |
| 121 | TTCAAGTGTCTATCTCTTGGCAGATCAGAACTCAGGAGTGTCTTTTCTTTTAATGCTTAATG | 180 |
| ib | | |
| 181 | GGAAARGACAGAAACCAACGACACTGGTGTGAAACCTTGCTATGFGTATAAAGATAAAGC | 240 |
| iy | | |
| 181 | GGAAARGACAGAAACCAACGACACTGGTGTGAAACCTTGCTATGFGTATAAAGATAAAGC | 240 |
| ib | | |
| 241 | GGCAGATTGTTTTGCTACTCGTGAAGAATAATTTCTGGTCCATTGGAATAGTGAAGCAAGG | 300 |
| iy | | |
| 241 | GGCAGATTGTTTTGCTACTCGTGAAGAATAATTTCTGGTCCATTGGAATAGTGAAGCAAGG | 300 |
| ib | | |
| 301 | TTGTTGGCTGGATGATATCAACTGCTATGACAGACTGATTGTGTNGAAAAAAGACAG | 360 |
| iy | | |
| 301 | TTGTTGGCTGGATGATATCAACTGCTATGACAGACTGATTGTGTATAAAAAAAGACAG | 360 |
| ib | | |
| 361 | CCCTGAAGTGTACTTTTGTGTGTGAGGGCAATATGTGTAAATGAAAGTTCTCTTATTT | 420 |
| iy | | |
| 361 | CCCTGAAGTGTACTTTTGTGTGTGAGGGCAATATGTGTAAATGAAAGTTCTCTTATTT | 420 |
| ib | | |
| 421 | TCGGAGATGGAAGTCACACAGCCCACTTCAATCCTGTTACACGAGGCCACCTATTAT | 480 |
| iy | | |
| 421 | TCGGAGATGGAAGTCACACAGCCCACTTCAATCCTGTTACACGAGGCCACCTATTAT | 480 |
| ib | | |
| 481 | CAACATTCTGCTGTAATTCCTGGTACCACTAATGTTAAATTGAGGAATGTCTATTGTGC | 540 |
| iy | | |
| 481 | CAACATTCTGCTGTAATTCCTGGTACCACTAATGTTAAATTGAGGAATGTCTATTGTGC | 540 |
| ib | | |
| 541 | ATTTTGGGTGACAGACATCAAGAATGGCCCTACCTCTGTACTTGTTCCTACTCAAGA | 600 |
| iy | | |
| 541 | ATTTTGGGTGACAGACATCAAGAATGGCCCTACCTCTGTACTTGTTCCTACTCAAGA | 600 |
| ib | | |
| 601 | CCCGAGACACCCCACTTCCCACTACTAGGGTTGAAGCCATTGCAGCTGTTAGAAGT | 660 |
| iy | | |
| 601 | CCCGAGACACCCCACTTCCCACTACTAGGGTTGAAGCCATTGCAGCTGTTAGAAGT | 660 |
| ib | | |
| 661 | GAAAGCAAGGGAAGATTGGTTGTCTGTGAAAGCCAGTTGCTCAATGAATATGTGC | 720 |
| iy | | |
| 661 | GAAAGCAAGGGAAGATTGGTTGTCTGTGAAAGCCAGTTGCTCAATGAATATGTGC | 720 |
| ib | | |
| 721 | TGCTCAAAATATTTCCAAATACAGACAAAACAGTCTGGCAGATGAATATGAAGTCTATAG | 780 |
| iy | | |
| 721 | TGCTCAAAATATTTCCAAATACAGACAAAACAGTCTGGCAGATGAATATGAAGTCTATAG | 780 |
| ib | | |
| 781 | TCTACTCGAATGAAGCATGAACAATCTACAGTTTCATTGGTGCAGAGAAAAGGGCAC | 840 |
| iy | | |
| 781 | TCTACTCGAATGAAGCATGAACAATCTACAGTTTCATTGGTGCAGAGAAAAGGGCAC | 840 |
| ib | | |
| 841 | CAGTGTGGATGTGGACCTGTGGCTAATCAGACATTTTCATGAAAAGGGCTCACTGTGAGA | 900 |
| iy | | |
| 841 | CAGTGTGGATGTGGACCTGTGGCTAATCAGACATTTTCATGAAAAGGGCTCACTGTGAGA | 900 |
| ib | | |
| 901 | CTTTCTTAAGGCTTAATGTGGTCTCTTTGGAATCAACCTTTGTCATATTCGAGAAACCAATGGC | 960 |
| iy | | |
| 901 | CTTTCTTAAGGCTTAATGTGGTCTCTTTGGAATCAACCTTTGTCATATTCGAGAAACCAATGGC | 960 |
| ib | | |
| 961 | TAGAGGATTTGGCATATTTACATGAGATATACCTGGCTTAAAGATGGCCACAGGCTGC | 1020 |
| iy | | |
| 961 | TAGAGGATTTGGCATATTTACATGAGATATACCTGGCTTAAAGATGGCCACAGGCTGC | 1020 |
| ib | | |
| 1021 | AATCTCTCACAGGACATCAAAAGTAAATGTGCTGTTGAAAAACAATCTGACAGCTTG | 1080 |
| iy | | |
| 1021 | AATCTCTCACAGGACATCAAAAGTAAATGTGCTGTTGAAAAACAATCTGACAGCTTG | 1080 |
| ib | | |
| 1081 | CATTGCTGACTTTTGGGTTGGCCCTTAAAGTTCGAGGCTGGCAAGTCTGCAAGTCAACCCCA | 1140 |
| iy | | |
| 1081 | CATTGCTGACTTTTGGGTTGGCCCTTAAAGTTCGAGGCTGGCAAGTCTGCAAGTCAACCCCA | 1140 |
| ib | | |
| 1141 | TGGGCAAGTTGGTACCCGGAGGTATATGGCTCCAGAGGTGTGTGGAGGGGTGCTATAAAGTT | 1200 |
| iy | | |

| | | | | |
|---|---|------|---|------|
| D | b | 1141 | TGGG CAGGT TGGTATCCCG GAGGTATATGGCTCAGAGGTGTGTGAGGGTGCTATAA ACTT | 1200 |
| Q | y | 1201 | CCAAGGACGCATTTC TGAGGATAGATATGTACGCCATGGGATTAGTCC TATGGGA ATT | 1260 |
| D | b | 1201 | CCAAGGACGCATTTC TGAGGATAGATATGTACGCCATGGGATTAGTCC TATGGGA ATT | 1260 |
| Q | y | 1261 | GGCTTCTCGTTGCACTGCTGCAGATGGACCCGTAGATGAGTACATGTTACCATT TTGAGGA | 1320 |
| D | b | 1261 | GGCTTCTCGTTGCACTGCTGCAGATGGACCCGTAGATGAGTACATGTTACCATT TTGAGGA | 1320 |
| Q | y | 1321 | AGAAATTGGCCAGCATCCATCTCTTGAAGATATGCAGGAAGTTGTTGTGCATATAAAAA AA | 1380 |
| D | b | 1321 | AGAAATTGGCCAGCATCCATCTCTTGAAGATATGCAGGAAGTTGTTGTGCATATAAAAA AA | 1380 |
| Q | y | 1391 | GAGGCTGTTTTAAGAGATTATTGGCAGAAAACAATGCAGGAATGSCAATGCTCTGTGA AAC | 1440 |
| D | b | 1391 | GAGGCTGTTTTAAGAGATTATTGGCAGAAAACAATGCAGGAATGSCAATGCTCTGTGA AAC | 1440 |
| Q | y | 1441 | GATAGAANAATGTTGGGATCATGATCCAGAACCCAGGTTATCAGCTGGATGTTAGGTGA | 1500 |
| D | b | 1441 | GATAGAANAATGTTGGGATCATGATCCAGAACCCAGGTTATCAGCTGGATGTTAGGTGA | 1500 |
| Q | y | 1501 | AAGAATTACTCAGATGCAAAAGCTAAACAATATCATTACTACAGAGGACATTGTAACAGT | 1560 |
| D | b | 1501 | AAGAATTACTCAGATGCAAAAGCTAAACAATATCATTACTACAGAGGACATTGTAACAGT | 1560 |
| Q | y | 1561 | GGTCACAATGGTGCACAAATGTTGACTTTCCTCCCAAAGAAATCTAGTCTATGATGGTGA | 1620 |
| D | b | 1561 | GGTCACAATGGTGCACAAATGTTGACTTTCCTCCCAAAGAAATCTAGTCTATGATGGTGA | 1620 |
| Q | y | 1621 | CCGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGTGCTTAAAGTAAAGGAAAGTGC | 1680 |
| D | b | 1621 | CCGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGTGCTTAAAGTAAAGGAAAGTGC | 1680 |
| Q | y | 1681 | TTAGTTGATTTTCTGTGTGAATGAGTAGGATGCCCTCCAGGACATGTACGACAGCAGCCC | 1740 |
| D | b | 1681 | TTAGTTGATTTTCTGTGTGAATGAGTAGGATGCCCTCCAGGACATGTACGACAGCAGCCC | 1740 |
| Q | y | 1741 | CTTGTGGAAAGCATGGATCTGGGAGATGGATCTGGGAAACTTACTGCACTGCTCAGCA | 1800 |
| D | b | 1741 | CTTGTGGAAAGCATGGATCTGGGAGATGGATCTGGGAAACTTACTGCACTGCTCAGCA | 1800 |
| Q | y | 1801 | CAGATATGAAGAGGAGTCTAAGGGGAAAAGCTGCAAACTGTAAAGAACTTCTGAAAATGTA | 1860 |
| D | b | 1801 | CAGATATGAAGAGGAGTCTAAGGGGAAAAGCTGCAAACTGTAAAGAACTTCTGAAAATGTA | 1860 |
| Q | y | 1861 | CTCGAAGAAATGCGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTT | 1920 |
| D | b | 1861 | CTCGAAGAAATGCGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTT | 1920 |
| Q | y | 1921 | GCAAACTGCACATCAGATTTCCTTAATGTCGTGTCAGAAAGACACTAATTCCTTAAATGAACT | 1980 |
| D | b | 1921 | GCAAACTGCACATCAGATTTCCTTAATGTCGTGTCAGAAAGACACTAATTCCTTAAATGAACT | 1980 |
| Q | y | 1981 | ACTGCTATTTTTTTTAAATGAAAACCTTTTCATTTCAGATTTTAAAGGGTAACTTTTT | 2040 |
| D | b | 1981 | ACTGCTATTTTTTTTAAATGAAAACCTTTTCATTTCAGATTTTAAAGGGTAACTTTTT | 2040 |
| Q | y | 2041 | ATTGCAATTTGCTGTGTTCTATAAATGACTATTGTAATGCAACATGCACACAGCTTGTG | 2100 |
| D | b | 2041 | ATTGCAATTTGCTGTGTTCTATAAATGACTATTGTAATGCAACATGCACACAGCTTGTG | 2100 |
| Q | y | 2101 | AATGTGTAGTGTGCTGCTTCTGTGTACATAGTGCATCAAAGTGGGTACAGTAAAGAGG | 2160 |
| D | b | 2101 | AATGTGTAGTGTGCTGCTTCTGTGTACATAGTGCATCAAAGTGGGTACAGTAAAGAGG | 2160 |
| Q | y | 2161 | CTTCCAAGCATTTACTTTAAACCTCCCTCAACAAGGTATACCTCAGTTCCACGGTTGTTAAA | 2220 |
| D | b | 2161 | CTTCCAAGCATTTACTTTAAACCTCCCTCAACAAGGTATACCTCAGTTCCACGGTTGTTAAA | 2220 |
| Q | y | 2221 | TTATATAAATTGAAAACACTTAAAGAAATTTGAAATAAATCAGTCCATGTTTTTATTAACAAGGT | 2280 |

Db 2221 TTATATAAATGAAACACATAACAGAAATTTGAATAAATCATGCTCATGTTTATATAACAAGGT 2280
Qy 2281 TAATTAACAATTCACCTGCTGTTATTTAAGAAAAAATGGTAAGCTATGCTTAGTGCCAATAG 2340
Db 2281 TAATTAACAATTCACCTGCTGTTATTTAAGAAAAAATGGTAAGCTATGCTTAGTGCCAATAG 2340
Qy 2341 TAAGTGGCTATTGTAAGACAGGTTTGTAGCTTTCTTCTACTGGCTGTGTAATTTAGGGA 2400
Db 2341 TAAGTGGCTATTGTAAGACAGGTTTGTAGCTTTCTTCTACTGGCTGTGTAATTTAGGGA 2400
Qy 2401 AAACAAGTCTCTCTTTGAAATGGAAGAAATATGTTGTACACCTACCCCTACTATAT 2460
Db 2401 AAACAAGTCTCTCTTTGAAATGGAAGAAATATGTTGTACACCTACCCCTACTATAT 2460
Qy 2461 ATCAAGGTCCTCAAAATATTTCTTTCAATTTCAAGACAGACATTTGAAACCCCTAAATTA 2520
Db 2461 ATCAAGGTCCTCAAAATATTTCTTTCAATTTCAAGACAGACATTTGAAACCCCTAAATTA 2520
Qy 2521 CAAGCCAGTAGAAGAAAGCTAAACACACGCTTTACAAATAGCC 2563
Db 2521 CAAGCCAGTAGAAGAAAGCTAAACACACGCTTTACAAATAGCC 2563

RESULT 6
AAC85297
ID AAC85297 standard; cDNA; 2563 BP.
XX
AC AAC85297;
DT 29-MAR-2001 (first entry)
XX
DE Mouse activin receptor cDNA.
XX
KW Mouse; human; rat; Xenopus; activin receptor; ligand binding domain;
KW transmembrane domain; receptor domain; serine kinase; TGF-beta;
KW transforming growth factor-beta; carcinogenesis; cancer; wound healing;
KW ss.
XX
OS Mus musculus.

Key Location/Qualifiers
FH 71..1612
CDS /*tag= a
FT /*product= "Activin receptor"
FT 71..127
FT /*tag= b
PN
PN US6162896-A.
PD 19-DEC-2000.
XX
PF 07-JUN-1995; 95US-00476123.
XX
PR 10-MAY-1991; 91US-00698709.
PR 08-OCT-1991; 91US-00773229.
PR 08-MAY-1992; 92US-00880220.
PR 02-SEP-1994; 94US-00300584.
XX
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Tsuchida K, Vale WW, Mathews LW;
XX
XX WPI; 2001-090408/10.
DR P-PSDB; AAB47025.
XX

Novel vertebrate activin receptor having extracellular ligand binding
domain, transmembrane domain and intracellular serine/threonine kinase
PT domain is useful for diagnosing and treating e.g. carcinogenesis, wound
PT healing.
XX
PS Claim 1; Col 21-26; 33pp; English.
XX
XX This sequence encodes a mouse-derived activin receptor. The activin

CC receptors of the invention comprise three distinct domains: an
CC extracellular, ligand binding domain, a hydrophobic, transmembrane domain
CC and an intracellular, receptor domain having serine kinase-like activity.
CC These proteins have binding affinity for at least one member of the
CC activin/TGF-beta superfamily of polypeptide growth factors. The activin
CC receptor proteins of the invention can be employed for a variety of
CC therapeutic uses, e.g. to block receptors. The presence of the soluble
CC proteins will compete with functional ligand for the receptor, preventing
CC the formation of a functional receptor-ligand complex, thereby blocking
CC the normal regulatory action of the complex. The receptor proteins are
CC useful for the diagnosis and therapeutic management of carcinogenesis,
CC wound healing, disorders of the immune, reproductive, or central nervous
CC systems
XX

SQ Sequence 2563 BP; 775 A; 483 C; 585 G; 720 T; 0 U; 0 Other;
Query Match 99.7%; Score 2556.4; DB 4; Length 2563;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2557; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCGAGGAAGACCCAGGAACTGGATATCTAGCGAGAACTTCCTACGGCTTCCTCGGCG 60
Db 1 CTCGAGGAAGACCCAGGAACTGGATATCTAGCGAGAACTTCCTACGGCTTCCTCGGCG 60
Qy 61 CCTCGGAAATGGGAGCTCTGCAAAAGTTGGCGTTGGCGCTTTCTTATCTCTGCTC 120
Db 61 CCTCGGAAATGGGAGCTCTGCAAAAGTTGGCGTTGGCGCTTTCTTATCTCTGCTC 120
Qy 121 TTCAGTGTCTACTTGGGAGATCAGAAACTCAGAGTGTCTTTCTTAATGCTAAATG 180
Db 121 TTCAGTGTCTACTTGGGAGATCAGAAACTCAGAGTGTCTTTCTTAATGCTAAATG 180
Qy 181 GGAAGAGCAGAACCAACAGACTGGTGTGAACCTTGTATGTGTAAGATAAAGC 240
Db 181 GGAAGAGCAGAACCAACAGACTGGTGTGAACCTTGTATGTGTAAGATAAAGC 240
Qy 241 GCGACATTTGTTTCTACTTGGGAGAAATATTTCTGGTCCATTTGAAATAGTGAAGCAAG 300
Db 241 GCGACATTTGTTTCTACTTGGGAGAAATATTTCTGGTCCATTTGAAATAGTGAAGCAAG 300
Qy 301 TTGTTGCTGGATGATATCAACTCTATGACAGACTGATGTAAGAAAAAAGACAG 360
Db 301 TTGTTGCTGGATGATATCAACTCTATGACAGACTGATGTAAGAAAAAAGACAG 360
Qy 361 CCTGAGTGTACTTTTGTGCTGCTGAGGCAATATGTAATGAAAGTCTCTTATTT 420
Db 361 CCTGAGTGTACTTTTGTGCTGCTGAGGCAATATGTAATGAAAGTCTCTTATTT 420
Qy 421 TCCGGAGATGGAAGTCAACAGCCCACTTCAAATCCTGTTCACCGAAGCCCTTATTA 480
Db 421 TCCGGAGATGGAAGTCAACAGCCCACTTCAAATCCTGTTCACCGAAGCCCTTATTA 480
Qy 481 CAACATTTCTGTATTTCCTTGGTACCACTAATGTTAATTTGAGGAATTTGTTGTC 540
Db 481 CAACATTTCTGTATTTCCTTGGTACCACTAATGTTAATTTGAGGAATTTGTTGTC 540
Qy 541 ATTTTGGGTGTACAGATCAAGATGCGCTTACCTCTCTGTACTTGTTCCTACTCAAGA 600
Db 541 ATTTTGGGTGTACAGATCAAGATGCGCTTACCTCTCTGTACTTGTTCCTACTCAAGA 600
Qy 601 CCCAGGACCAACCCCTTCCCATTTACTAGGTTGAGCCATTTGCAGCTGTAGAAAGT 660
Db 601 CCCAGGACCAACCCCTTCCCATTTACTAGGTTGAGCCATTTGCAGCTGTAGAAAGT 660
Qy 661 GAAAGCAAGGGGAAGATTTGTTGTCTGGAAGCCCAAGTTGCTCAATGAATATGTGC 720
Db 661 GAAAGCAAGGGGAAGATTTGTTGTCTGGAAGCCCAAGTTGCTCAATGAATATGTGC 720
Qy 721 TGTCAAAATATTTCCAAATACAGGACAAACAGTCTTGGGCAAGATGAATATGAGTCTATAG 780
Db 721 TGTCAAAATATTTCCAAATACAGGACAAACAGTCTTGGCAGAAATGAATATGAGTCTATAG 780
Qy 781 TCTACCTGGAAATGAAGCATGAGACATACACTAGTTTCATTTGGTGCAGAGAAAGACAC 840

840 841 841 901 901 961 961 1021 1021 1081 1081 1141 1141 1201 1201 1261 1261 1321 1321 1381 1381 1441 1441 1501 1501 1561 1561 1621 1621 1681 1681 1741 1741 1801 1801 1861

TCCTACCTGGAAATGAAGCATGAGACATCTACAGTTTCATGGTGCAGAAAGAGGCAC
CAGTGTGGATGTGACCTGTGGCTAATCACAGCACTTTCATGAAAGGGCTCACTGTGAGA
CAGTGTGGATGTGACCTGTGGCTAATCACAGCACTTTCATGAAAGGGCTCACTGTGAGA
CTTTCTTAAGGCTAATGTGGTCTCTCTGGAAAGCACTTTGTGCATATTCAGAAACCATGGC
CTTTCTTAAGGCTAATGTGGTCTCTCTGGAAAGCACTTTGTGCATATTCAGAAACCATGGC
TAGAGGATGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAGGCTGC
TAGAGGATGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAGGCTGC
AATCTCTCAGAGGACATCAAAAGTAAAGATGTGCTGTGAAAGCAATCTGCAGCTTG
AATCTCTCAGAGGACATCAAAAGTAAAGATGTGCTGTGAAAGCAATCTGCAGCTTG
CAITGTCTGACTTTGGTGTGGCTTAAAGTTCAGAGCTGGCAAGTCTGCAGGTGACACCCA
CAITGTCTGACTTTGGTGTGGCTTAAAGTTCAGAGCTGGCAAGTCTGCAGGTGACACCCA
TGAGGAGGTGGTACCCGAGGATATATGGCTCCAGAGGTGTGGAGGGTGCATATAACTT
TGAGGAGGTGGTACCCGAGGATATATGGCTCCAGAGGTGTGGAGGGTGCATATAACTT
CCAAAGGAGCGCATTTCTGAGGATAGATATGACGCCATGGGATAGTCTCTATGGGAAT
CCAAAGGAGCGCATTTCTGAGGATAGATATGACGCCATGGGATAGTCTCTATGGGAAT
GGCTTCCTGTTGACTGCTGAGATGGAGCCCTAGATGAGTACATGTTACCATTTGAGGA
GGCTTCCTGTTGACTGCTGAGATGGAGCCCTAGATGAGTACATGTTACCATTTGAGGA
AGAAATTTGGCCAGCATCCATCTCTTGAAGATATGAGGAGGAGTGTGTGATATAAAAAA
AGAAATTTGGCCAGCATCCATCTCTTGAAGATATGAGGAGGAGTGTGTGATATAAAAAA
GAGGCTGTTTGAAGATATTTGCGAGAAATATGCGAGGAATGCGATCTGTGGAAC
GAGGCTGTTTGAAGATATTTGCGAGAAATATGCGAGGAATGCGATCTGTGGAAC
GATAGAGATGTTGGGATCATGATGCGAGGAGGAGTATCAGCTGGATGTGTAGGTGA
GATAGAGATGTTGGGATCATGATGCGAGGAGGAGTATCAGCTGGATGTGTAGGTGA
AAGAAATTAATCAGATGCAAGCACTAACCAATATCTACTACAGAGGACATTTCTAACAGT
AAGAAATTAATCAGATGCAAGCACTAACCAATATCTACTACAGAGGACATTTCTAACAGT
GGTCACAATGGTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGGTGCA
GGTCACAATGGTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGGTGCA
CCGTCTGTACACATGAGGACTGGGACTCTGAATCTGAGAGCTGTAAGTAAAGAAAGTGC
CCGTCTGTACACATGAGGACTGGGACTCTGAATCTGAGAGCTGTAAGTAAAGAAAGTGC
TTAGTGTATTTCTGTGGAATGATGAGTGGCTCCAGGACATGTACGCAAGGAGCC
TTAGTGTATTTCTGTGGAATGATGAGTGGCTCCAGGACATGTACGCAAGGAGCC
CTTGTGAAAGCATGGATCTGGGAGATGGATCTGGGAAATCTACTGCATCTGTGAGCA
CTTGTGAAAGCATGGATCTGGGAGATGGATCTGGGAAATCTACTGCATCTGTGAGCA
CAGATATGAAGAGGAGTCTAAGGGGAAAAAGCTGCAAACTGTGAAAGAACTTCTGAAAAATGA
CAGATATGAAGAGGAGTCTAAGGGGAAAAAGCTGCAAACTGTGAAAGAACTTCTGAAAAATGA
CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAATCAAGTATTT

1861 1921 1921 1981 1981 2041 2041 2101 2101 2161 2161 2221 2221 2281 2281 2341 2341 2401 2401 2461 2461 2521 2521 2563 2563

CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGCTAATCAAGTATTT
GCAAAACCTGACATCAGATTTCTTAATGTCTGTGAGAGACACTAATTCCTTAATGAAC
GCAAAACCTGACATCAGATTTCTTAATGTCTGTGAGAGACACTAATTCCTTAATGAAC
ACTGCTATTTTCTTAATGAAGAACTTTTCAATTCAGATTTTAAAGAGGTAACCTTTT
ACTGCTATTTTCTTAATGAAGAACTTTTCAATTCAGATTTTAAAGAGGTAACCTTTT
ATTGCATTTGCTGTGTTTCTTAATGAAGAACTTTTCAATTCAGATTTTAAAGAGGTA
ATTGCATTTGCTGTGTTTCTTAATGAAGAACTTTTCAATTCAGATTTTAAAGAGGTA
AATGTGTAGTGTGCTGTGTTTCTGTGTACATAGTCAATCAAGTGGGGTACAGTAAAGAGG
AATGTGTAGTGTGCTGTGTTTCTGTGTACATAGTCAATCAAGTGGGGTACAGTAAAGAGG
CTTCCAGCATTAACCTTAACTCCCTCAACAGGATATACCTCAGTTCACGGTGTGTTAAA
TTATAAATTTGAAACACATTAACAGAAATTTGAATTAATCAGTCCATGTTTATTAACAGG
TTATAAATTTGAAACACATTAACAGAAATTTGAATTAATCAGTCCATGTTTATTAACAGG
TAATTTACAAATTTCACTGTGTTTAAAGAAATTTGTAAGTATGCTTGTAGTGCAATAG
TAATTTACAAATTTCACTGTGTTTAAAGAAATTTGTAAGTATGCTTGTAGTGCAATAG
TAAGTGGCTATTTGTAAGAGCAGTGTGTTTCTTCTACTGGCTGTGTAATTTAGGGA
TAAGTGGCTATTTGTAAGAGCAGTGTGTTTCTTCTACTGGCTGTGTAATTTAGGGA
AAACAGTGTGCTCTTTGAAATGGAAGAAATATGCTGTACCCCTACCCCTACCTATAT
AAACAGTGTGCTCTTTGAAATGGAAGAAATATGCTGTACCCCTACCCCTACCTATAT
ATCAAGTGTGCTCTTTTCCATTTTCAAGGACAGCACTTTGAAACCCCTAAATTA
ATCAAGTGTGCTCTTTTCCATTTTCAAGGACAGCACTTTGAAACCCCTAAATTA
CAAGCAGTGTGAGAAAGAAAGCTTAAACACAGCTTTTACAAATAGCC 2563
CAAGCAGTGTGAGAAAGAAAGCTTAAACACAGCTTTTACAAATAGCC 2563

RESULT 7
AAK98720
ID AAK98720 standard; cDNA; 2563 BP.
XX AAK98720;
AC AC
XX AC
DT 07-MAY-2002 (first entry)
XX cDNA encoding a mouse-derived activin receptor.
DE Mouse; activin receptor; cloning; recombinant; TGF-beta; ligand-binding;
XX superfamily; trans-membrane; receptor domain; serine kinase; diagnosis;
XX therapeutic management; carcinogenesis; wound healing; protein therapy;
KW immune; reproductive; central nervous system; activin-dependent tumour;
KW brain neuron; abortion; livestock; twinning; agonist; cytostatic;
KW wound healing; transplant organ rejection; vulnery; immunosuppressive;
transforming growth factor-beta; ss.
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
PH 1..70
5'UTR /tag= a
FT 71..1612
FT CDS /tag= b
FT /product= "Mouse-derived activin receptor protein"

| | | | |
|----|---|------------|--|
| FT | 3'UTR | 1613..2563 | |
| FT | | /*tag= c | |
| XX | US2001039036-A1. | | |
| XX | 08-NOV-2001. | | |
| XX | 19-DEC-2000; 2000US-00742684. | | |
| XX | 10-MAY-1991; 91US-00698709. | | |
| PR | 09-OCT-1991; 91US-00773229. | | |
| PR | 08-MAY-1992; 92US-00880220. | | |
| PR | 02-SEP-1994; 94US-00300584. | | |
| PR | 07-JUN-1995; 95US-00476123. | | |
| XX | (SALK) SALK INST BIOLOGICAL STUDIES. | | |
| XX | Mathews LS, Vale WW, Tsuchida K; | | |
| PI | WPI: 2002-040721/05. | | |
| XX | P-PSDB; AA014118. | | |
| DR | | | |
| DR | | | |
| XX | New receptor proteins having an extracellular ligand-binding domain, a | | |
| PT | hydrophobic trans-membrane domain, and an intracellular receptor domain, | | |
| PT | useful for diagnosing or treating carcinogenesis, wound healing or immune | | |
| PT | disorders. | | |
| XX | | | |
| XX | Example 11; Page 12-15; 33pp; English. | | |
| PS | | | |
| XX | This polynucleotide sequence represents the cDNA encoding a mouse-derived | | |
| CC | activin receptor. The invention relates to cloning and recombinant | | |
| CC | production of receptor(s) of the activin/TGF-beta (transforming growth | | |
| CC | factor-beta) superfamily. The invention has identified and characterised | | |
| CC | members of a new superfamily of receptor proteins which comprise three | | |
| CC | distinct domains: an extracellular, ligand-binding domain, a hydrophobic, | | |
| CC | trans-membrane domain, and an intracellular, receptor domain having | | |
| CC | serine kinase-like activity. The receptor proteins and antibodies to | | |
| CC | these proteins are useful in the diagnosis and therapeutic management of | | |
| CC | carcinogenesis, wound healing, disorders of the immune, reproductive, or | | |
| CC | central nervous systems. The receptor proteins of the invention can be | | |
| CC | used in protein therapy. These may further be used to diagnose or treat | | |
| CC | activin-dependent tumours, enhance the survival of brain neurons, induce | | |
| CC | abortion in livestock and other domesticated animals, and induce twinning | | |
| CC | in livestock and other domesticated animals. The DNAs are useful as | | |
| CC | probes for identifying additional members of the superfamily of receptor | | |
| CC | proteins, and as coding sequences which can be used for the recombinant | | |
| CC | expression of the receptor proteins. Agonists for TGF-beta specific | | |
| CC | receptors can be used to stimulate wound healing, to suppress growth of | | |
| CC | TGF-beta sensitive tumours, or to suppress immune response (thus prevent | | |
| CC | rejection of transplant organ). The receptor proteins of the invention | | |
| CC | have cytostatic, vulnerary, and immunosuppressive activity | | |
| XX | | | |
| SQ | Sequence 2563 BP; 775 A; 483 C; 585 G; 720 T; 0 U; 0 Other; | | |
| | Query Match 99.7%; Score 2556.4; DB 6; Length 2563; | | |
| | Best Local Similarity 99.8%; Pred. No. 0; | | |
| | Matches 2557; Conservative 2; Mismatches 4; Indels 0; Gaps 0; | | |
| QY | 1 CTCGGAGAGACCCAGGAACTGGATATCAGGAGAACTTCCTACGGCTTCCTCGGGG 60 | | |
| DB | 1 CTCGGAGAGACCCAGGAACTGGATATCAGGAGAACTTCCTACGGCTTCCTCGGGG 60 | | |
| QY | 61 CCTCGGAAAATGGAGCTGCTGCAAGCTTGGCGTTCGCGCTTCTTATCTCTTGCTC 120 | | |
| DB | 61 CCTCGGAAAATGGAGCTGCTGCAAGCTTGGCGTTCGCGCTTCTTATCTCTTGCTC 120 | | |
| QY | 121 TTCAGGTGCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTTCTTTAATGCTAATTG 180 | | |
| DB | 121 TTCAGGTGCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTTCTTTAATGCTAATTG 180 | | |
| QY | 181 GGAAGAAGCAGAACCAACAGACTGGTGTGAACTTGTCTATGGTGATTAAGATAAAGC 240 | | |
| DB | 181 GGAAGAAGCAGAACCAACAGACTGGTGTGAACTTGTCTATGGTGATTAAGATAAAGC 240 | | |

| | | | |
|----|--|--|--|
| QY | 241 GGCACATTTGTTTGTCTACCTGGGAAGATAATTTCTGGTTCCATTGAAATAGTGAACAAGG 300 | | |
| DB | 241 GGCACATTTGTTTGTCTACCTGGGAAGATAATTTCTGGTTCCATTGAAATAGTGAACAAGG 300 | | |
| QY | 301 TTGTTGGCTGGATGATATCAACTGCTATGACAGGACTGATGTGTGTAAGAAAAAGACAG 360 | | |
| DB | 301 TTGTTGGCTGGATGATATCAACTGCTATGACAGGACTGATGTGTGTAAGAAAAAGACAG 360 | | |
| QY | 361 CCTGGAAGTGTACTTTTGTGCTGTGAGGCAATATGTAATGAAAAAGTTCTCTTATTT 420 | | |
| DB | 361 CCTGGAAGTGTACTTTTGTGCTGTGAGGCAATATGTAATGAAAAAGTTCTCTTATTT 420 | | |
| QY | 421 TCCGAGATGGAAGTACACAGGCCACTTCAAATCTCTGTTACACCGAAGCCACCTATTTA 480 | | |
| DB | 421 TCCGAGATGGAAGTACACAGGCCACTTCAAATCTCTGTTACACCGAAGCCACCTATTTA 480 | | |
| QY | 481 CAACATCTGCTATTTCTTGTGTACCACTTAATGTTAATTCGACGAATGTCTATTGTC 540 | | |
| DB | 481 CAACATCTGCTATTTCTTGTGTACCACTTAATGTTAATTCGACGAATGTCTATTGTC 540 | | |
| QY | 541 ATTTTGGGTGTACAGACATCAAGATGGCTTCCCTCTCTACTTGTCTTACTCAAGA 600 | | |
| DB | 541 ATTTTGGGTGTACAGACATCAAGATGGCTTCCCTCTCTACTTGTCTTACTCAAGA 600 | | |
| QY | 601 CCCAGGACCAACCCCTTCCCTTACTAGGGTTGAAGCCATTGCGAGCTGTAGAACT 660 | | |
| DB | 601 CCCAGGACCAACCCCTTCCCTTACTAGGGTTGAAGCCATTGCGAGCTGTAGAACT 660 | | |
| QY | 661 GAAAGCAAGGGGAAGATTGTTGTTGTCTGGAAGCCCGAGTTGCTCAATGAATATGTGC 720 | | |
| DB | 661 GAAAGCAAGGGGAAGATTGTTGTTGTCTGGAAGCCCGAGTTGCTCAATGAATATGTGC 720 | | |
| QY | 721 TGTCAAAATATTTCCAAATACAGCAAAACAGCTCTGGCAGAAATGAATGAAGTCTATAG 780 | | |
| DB | 721 TGTCAAAATATTTCCAAATACAGCAAAACAGCTCTGGCAGAAATGAATGAAGTCTATAG 780 | | |
| QY | 781 TCTACCTGGATGAAGCATGAGACATACAGTTCAATTTGTTGTCAGAGAAAGAGGCAC 840 | | |
| DB | 781 TCTACCTGGATGAAGCATGAGACATACAGTTCAATTTGTTGTCAGAGAAAGAGGCAC 840 | | |
| QY | 841 CAGTGTGGATGTGGACCTGTGGCTTAATCAGACATTTTCATGAAAAAGGCTCTACTGTGCA 900 | | |
| DB | 841 CAGTGTGGATGTGGACCTGTGGCTTAATCAGACATTTTCATGAAAAAGGCTCTACTGTGCA 900 | | |
| QY | 901 CTTTCTTAAAGCTAATGTGGTCTCTTGGAAATCARTTTTGTCTATTTGCAAGAACCAATGGC 960 | | |
| DB | 901 CTTTCTTAAAGCTAATGTGGTCTCTTGGAAATCARTTTTGTCTATTTGCAAGAACCAATGGC 960 | | |
| QY | 961 TAGAGGATTTGGCATATTTACATGAGGATATACCTGGCTTAAAGATGCGCCAGAGCTGC 1020 | | |
| DB | 961 TAGAGGATTTGGCATATTTACATGAGGATATACCTGGCTTAAAGATGCGCCAGAGCTGC 1020 | | |
| QY | 1021 AATCTCTCAGAGGACATCAAAAGTAAATAATGTGTTGAAAAACAATCTGCAGCTTG 1080 | | |
| DB | 1021 AATCTCTCAGAGGACATCAAAAGTAAATAATGTGTTGAAAAACAATCTGCAGCTTG 1080 | | |
| QY | 1081 CATTGCTGACTTTGGGTTGGCTTTAAAGTTGAGGCTGGCAAGTCTGCAAGTGCACACCA 1140 | | |
| DB | 1081 CATTGCTGACTTTGGGTTGGCTTTAAAGTTGAGGCTGGCAAGTCTGCAAGTGCACACCA 1140 | | |
| QY | 1141 TGGCAGGTTGGTACCCGGAGGTATATGGCTCCAGAGTGTGGAGGGTGTCTATAAATT 1200 | | |
| DB | 1141 TGGCAGGTTGGTACCCGGAGGTATATGGCTCCAGAGTGTGGAGGGTGTCTATAAATT 1200 | | |
| QY | 1201 CMAAGGAGCGCATTTCTGAGGATAGATATGTACGCCATGGGATAGTCTTATGGGAATT 1260 | | |
| DB | 1201 CMAAGGAGCGCATTTCTGAGGATAGATATGTACGCCATGGGATAGTCTTATGGGAATT 1260 | | |
| QY | 1261 GGCTTCTCGTTGCACTGCTGCAGATGGACCGTAGATGATGATCATGTTTACCATTGTGAGGA 1320 | | |
| DB | 1261 GGCTTCTCGTTGCACTGCTGCAGATGGACCGTAGATGATGATCATGTTTACCATTGTGAGGA 1320 | | |

1321 AGAAATGGCCAGCATCCATCTCTTGAAGATATGCGAAGTGTGTGTCATAAAAAA 1380
1321 AGAAATGGCCAGCATCCATCTCTTGAAGATATGCGAAGTGTGTGTCATAAAAAA 1380
1381 GAGGCGCTGTTTAAAGAGATTATTGGCAGAAACATGCGAGGATGCAATGCTGTGGAAC 1440
1381 GAGGCGCTGTTTAAAGAGATTATTGGCAGAAACATGCGAGGATGCAATGCTGTGGAAC 1440
1441 GATAGAAGAAATGTTGGGATCATGATGAGAGCCAGGTTATCAGCTGGATGTGTAAGTGA 1500
1441 GATAGAAGAAATGTTGGGATCATGATGAGAGCCAGGTTATCAGCTGGATGTGTAAGTGA 1500
1501 AAGAATTACTCAGATGCAAAAGACTAAACAAATATCACTACTACAGAGACATTTAAGT 1560
1501 AAGAATTACTCAGATGCAAAAGACTAAACAAATATCACTACTACAGAGACATTTAAGT 1560
1561 GGTCAAAATGGTGACAAATGTTGACTTTCCTCCCAAGAAATCTAGTCTATGATGTTGCA 1620
1561 GGTCAAAATGGTGACAAATGTTGACTTTCCTCCCAAGAAATCTAGTCTATGATGTTGCA 1620
1621 CGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAAGTGC 1680
1621 CGTCTGTACACACTGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAAGTGC 1680
1681 TTAGTTGATTTCTGTGTAATGAGTAGGATGCGTCCAGGACATGTACGCAAGCAGGCC 1740
1681 TTAGTTGATTTCTGTGTAATGAGTAGGATGCGTCCAGGACATGTACGCAAGCAGGCC 1740
1741 CTGTGGAAGCAATGATCTGGGAGATGGATCTGGGAAACTTACTGCATCGTCTGCAGCA 1800
1741 CTGTGGAAGCAATGATCTGGGAGATGGATCTGGGAAACTTACTGCATCGTCTGCAGCA 1800
1801 CAGATATGAGAGGAGCTAAGGAGAAAGCTGCAAACTGTAAGAACTCTCAGAAATGTA 1860
1801 CAGATATGAGAGGAGCTAAGGAGAAAGCTGCAAACTGTAAGAACTCTCAGAAATGTA 1860
1861 CTGCAAGAAATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAATCAAGTATTT 1920
1861 CTGCAAGAAATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAATCAAGTATTT 1920
1921 GCAAACTGACATCAGATTTCTTAATGCTGTGCAAGACACATTAATCTTAAATGAAT 1980
1921 GCAAACTGACATCAGATTTCTTAATGCTGTGCAAGACACATTAATCTTAAATGAAT 1980
1981 ACTGCTATTTTTTAAATGAAACCTTTTCAATTCAGATTTTAAAGGGTAACTTTTT 2040
1981 ACTGCTATTTTTTAAATGAAACCTTTTCAATTCAGATTTTAAAGGGTAACTTTTT 2040
2041 ATTGCATTTGCTGTTTCTATTAATGACTATTCTAATGCTAATGCTAATGCTGTTG 2100
2041 ATTGCATTTGCTGTTTCTATTAATGACTATTCTAATGCTAATGCTAATGCTGTTG 2100
2101 AATGTGTAGTGTGCTGTTTCTGTGTACATAGTATCAATCAAAAGTGGGTACAGTAAGAGG 2160
2101 AATGTGTAGTGTGCTGTTTCTGTGTACATAGTATCAATCAAAAGTGGGTACAGTAAGAGG 2160
2161 CTTCCAGCAATTTACTTTAACTCCCTCAACAGGATATACCTCAGTTCCAGGTTGTAA 2220
2161 CTTCCAGCAATTTACTTTAACTCCCTCAACAGGATATACCTCAGTTCCAGGTTGTAA 2220
2221 TTATAAATGAAACACTAAGCAATTTGAATAATCAGTCCAGTGTATTAACAGGT 2280
2221 TTATAAATGAAACACTAAGCAATTTGAATAATCAGTCCAGTGTATTAACAGGT 2280
2281 TAATTAATAATTTCACTGTTTATTTAAGAAATGTTAAGTATCTAGTCCCAATAG 2340
2281 TAATTAATAATTTCACTGTTTATTTAAGAAATGTTAAGTATCTAGTCCCAATAG 2340
2341 TAAGTGCTATTTTAAAGCAGTGTGTTTACTGCTTTTCTTACTGGCTGTAAATAGGGA 2400
2341 TAAGTGCTATTTTAAAGCAGTGTGTTTACTGCTTTTCTTACTGGCTGTAAATAGGGA 2400
2401 AAACAAGTGTCTCTTTGAAATGGAAGAAATATGGTGTACCCCTACCCCCCACTTAT 2460

Db 2401 AAACAAGTGTCTCTTTGAAATGGAAGAAATATGGTGTACCCCTACCCCTACCTTAT 2460
QY 2461 ATCAAGTCCCAAAATATCTTTTCCATTTTCAAGACAGACACTTTGAAAACCCCTAAATTA 2520
Db 2461 ATCAAGTCCCAAAATATCTTTTCCATTTTCAAGACAGACACTTTGAAAACCCCTAAATTA 2520
QY 2521 CAAGCCAGTAGAAGAAAGCTAAACACGCTTTTACAAATAGCC 2563
Db 2521 CAAGCCAGTAGAAGAAAGCTAAACACGCTTTTACAAATAGCC 2563

RESULT 8
AAQ31910
ID AAQ31910 standard; cDNA; 2563 BP.
XX
AC AAQ31910;
XX
DT 25-MAR-2003 (revised)
DT 19-APR-1993 (first entry)
XX
DE Mouse activin receptor.
XX
KW Activin receptor; mouse; Xenopus; human; extracellular; ligand binding;
KW hydrophobic; trans-membrane; intracellular; receptor; domain;
KW serine kinase-like; activity; probe; superfamily; secretion signal;
KW golgi membrane; diagnosis; treatment; activin-dependent tumour; brain;
KW neuron; abortion; twinning; wound healing; TGF-beta; immune response;
KW liver regeneration; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 71..1612
FT /*tag= a
XX
PN WO9220793-A1.
XX
PD 26-NOV-1992.
XX
PF 08-MAY-1992; 92WO-US003825.
XX
PR 10-MAY-1991; 91US-00698709.
PR 09-OCT-1991; 91US-00773229.
XX
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Mathews LS, Vale WW;
XX
DR WPI: 1992-415771/50.
DR P-PSDB; AAR29581.
XX
PT New member of activin-transforming growth factor beta super-family - for
PT diagnosis and treatment of cancer and disorders of the immune,
PT reproductive or central nervous system.
XX
PS Disclosure; Page 42-45; 68pp; English.
XX
CC The sequences given in AAQ31910-12 encode activin receptors from mouse,
CC Xenopus and human respectively. Each of the proteins encoded by these
CC sequences comprise three distinct domains: an extracellular, ligand
CC binding domain, a hydrophobic, trans-membrane domain and an
CC intracellular, receptor domain having serine kinase-like activity. These
CC DNA sequences can be used as probes for the identification of additional
CC members of this superfamily of receptor molecules. The proteins may
CC further comprise a second hydrophobic domain at the amino terminal which
CC comprises a secretion signal sequence which promotes the intracellular
CC transport of the initially expressed receptor protein across the golgi
CC membrane. The receptor proteins can be used to develop agents for the
CC diagnosis and/or treatment of eg. activin-dependent tumours, for
CC enhancing the survival of brain neurons, for inducing abortion or
CC twinning in livestock, for stimulating wound healing, for suppression of
CC growth of TGF-beta sensitive tumours, for suppressing immune response,

CC for promoting liver regeneration and for stimulating some immune
CC responses. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2563 BP; 776 A; 485 C; 563 G; 719 T; 0 U; 0 Other;

| | | | | | |
|-----------------------|---|--------------|---------------|------------|--------------|
| Query Match | | 99.6%; | Score 2551.6; | DB 2; | Length 2563; |
| Best Local Similarity | | 99.6%; | Pred. No. 0; | | |
| Matches 2554; | | Conservative | 2; | Mismatches | 7; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| 1 | CTCCGAGGAAGACCCAGGAACTGGATATCTAGCGAGAACTTCTCTACGGCTTCTCCGGCG | 60 | | | |
| 1 | CTCCGAGGAGACCCAGGAACTGGATATCTAGCGAGAACTTCTCTACGGCTTCTCCGGCG | 60 | | | |
| 61 | CCTCGGAAAAATGGAGCTGCTGCAAAAGTTGGCGTTCGCCGCTTCTTTATCTCTGTCTC | 120 | | | |
| 61 | CCTCGGAAAAATGGAGCTGCTGCAAAAGTTGGCGTTCGCCGCTTCTTTATCTCTGTCTC | 120 | | | |
| 121 | TTTCAGGTGCTATCTTGGCAGATCAGAAACTCAGGAGTGTCTTTCTTTAATGCTAAATTG | 180 | | | |
| 121 | TTTCAGGTGCTATCTTGGCAGATCAGAAACTCAGGAGTGTCTTTCTTTAATGCTAAATTG | 180 | | | |
| 181 | GGAAARGACAGAACCAACAGACTGGTGTGAACCTTGCTATGGTGATAAAGATAAACG | 240 | | | |
| 181 | GGAAAGACAGAACCAACAGACTGGTGTGAACCTTGCTATGGTGATAAAGATAAACG | 240 | | | |
| 241 | GGCAGATTTGTTGTACTGGAAGAAATATTTCTGGTTCCATTGAAATAGTGAAGCAAGG | 300 | | | |
| 241 | GGCAGATTTGTTGTACTGGAAGAAATATTTCTGGTTCCATTGAAATAGTGAAGCAAGG | 300 | | | |
| 301 | TTGTTGGCTGGATGATATCAACTGCTATGACAGGACTGAATGTGTGCAAAAAAGACAG | 360 | | | |
| 301 | TTGTTGGCTGGATGATATCAACTGCTATGACAGGACTGAATGTATGAAAAAAGACAG | 360 | | | |
| 361 | CCCTGAAGTGTACTTTTGTGCTGTGAGGCAATATGTGTAATGAAAGTTCTCTTAATT | 420 | | | |
| 361 | CCCTGAAGTGTACTTTTGTGCTGTGAGGCAATATGTGTAATGAAAGTTCTCTTAATT | 420 | | | |
| 421 | TCGGAGATGGAAGTCAACAGCCCACTTCAAAATCCTGTTTACCCGAGCCACCTTAATTA | 480 | | | |
| 421 | TCGGAGATGGAAGTCAACAGCCCACTTCAAAATCCTGTTTACCCGAGCCACCTTAATTA | 480 | | | |
| 481 | CAACATTTCTGCTGATTTCTTGTGTACCACTAATGTAAATTCGAGGAATGTCTATTGTGC | 540 | | | |
| 481 | CAACATTTCTGCTGATTTCTTGTGTACCACTAATGTAAATTCGAGGAATGTCTATTGTGC | 540 | | | |
| 541 | ATTTGGGTGTACAGACATCAAGATGGCCCTACCTCCTGTACTTGTTCCTACTCAAGA | 600 | | | |
| 541 | ATTTGGGTGTACAGACATCAAGATGGCCCTACCTCCTGTACTTGTTCCTACTCAAGA | 600 | | | |
| 601 | CCAGGACACCCCACTTCCCAATTAAGGTTGAAGCCATTCAGCTGTTAGAAAT | 660 | | | |
| 601 | CCAGGACACCCCACTTCCCAATTAAGGTTGAAGCCATTCAGCTGTTAGAAAT | 660 | | | |
| 661 | GAAAGCAAGGGGAAGATTTGGTGTGCTGGAAGCCAGTGTCTCAATGAATATGTGGC | 720 | | | |
| 661 | GAAAGCAAGGGGAAGATTTGGTGTGCTGGAAGCCAGTGTCTCAATGAATATGTGGC | 720 | | | |
| 721 | TGTCAAAATATTTTCCAATACAGGACAAACAGTCTCTGGCAGAAATATGAAGTCTATAG | 780 | | | |
| 721 | TGTCAAAATATTTTCCAATACAGGACAAACAGTCTCTGGCAGAAATATGAAGTCTATAG | 780 | | | |
| 781 | TCTACTGGAATGAGCATGAGAAACATACTACAGTTCAATGGTGTGAGAAAAAGGGCAC | 840 | | | |
| 781 | TCTACTGGAATGAGCATGAGAAACATACTACAGTTCAATGGTGTGAGAAAAAGGGCAC | 840 | | | |
| 841 | CAGTGTGGAATGTGGAACCTGTGGCTAATACAGCAATTTCAATGAAGGGCTCACTGTGAGA | 900 | | | |
| 841 | CAGTGTGGAATGTGGAACCTGTGGCTAATACAGCAATTTCAATGAAGGGCTCACTGTGAGA | 900 | | | |
| 901 | CTTTCTTAAGGCTAATGTGGTCTCTTGGAAATCACTTTGTGATATGCAAAAAACATGGC | 960 | | | |
| 901 | CTTTCTTAAGGCTAATGTGGTCTCTTGGAAATCACTTTGTGATATGCAAAAAACATGGC | 960 | | | |

| | | | |
|----|------|---|------|
| QY | 961 | TAGAGGATTCGCATATTTTACATGAGGATATACCTGGCTTAAAGATGGCCACAGCCTGC | 1020 |
| DB | 961 | TAGAGGATTCGCATATTTTACATGAGGATATACCTGGCTTAAAGATGGCCACAGCCTGC | 1020 |
| QY | 1021 | AATCTCTCAGAGGACATCAAAAGTAAAGTGTCTGTGTAAGAAACATCTGACAGCTTG | 1080 |
| DB | 1021 | AATCTCTCAGAGGACATCAAAAGTAAAGTGTCTGTGTAAGAAACATCTGACAGCTTG | 1080 |
| QY | 1081 | CATTGTGACTTTTGGGTTGGCTTAAAGTTCAGGCTGGCAAGTCTGACAGGTGACACCCA | 1140 |
| DB | 1081 | CATTGTGACTTTTGGGTTGGCTTAAAGTTCAGGCTGGCAAGTCTGACAGGTGACACCCA | 1140 |
| QY | 1141 | TGGGCAAGTTGGTACCCGAGGATATATGGCTCCAGAGGTGTTGGAGGTTGCTATAAAGTT | 1200 |
| DB | 1141 | TGGGCAAGTTGGTACCCGAGGATATATGGCTCCAGAGGTGTTGGAGGTTGCTATAAAGTT | 1200 |
| QY | 1201 | CCAAAGGACGCAATTTCTGAGGATAGATATGTAAGTTCAGGCTGGCAAGTCTGAGGTTGGAAT | 1260 |
| DB | 1201 | CCAAAGGACGCAATTTCTGAGGATAGATATGTAAGTTCAGGCTGGCAAGTCTGAGGTTGGAAT | 1260 |
| QY | 1261 | GGCTTCTCGTTGCACTGCTCAGATGGAACCGTGAATGAGTACATGTTACCATTTTGAGGA | 1320 |
| DB | 1261 | GGCTTCTCGTTGCACTGCTCAGATGGAACCGTGAATGAGTACATGTTACCATTTTGAGGA | 1320 |
| QY | 1321 | AGAAATTTGGCCAGCATCCATCTCTTGAAGATATGCGAAGTTGTTGTGCTATAAAAAAA | 1380 |
| DB | 1321 | AGAAATTTGGCCAGCATCCATCTCTTGAAGATATGCGAAGTTGTTGTGCTATAAAAAAA | 1380 |
| QY | 1381 | GAGGCTGTTTTAAGAGATTATTGGCAGAAACATGCGAGAAATGCAATGCTCTGTGAAC | 1440 |
| DB | 1381 | GAGGCTGTTTTAAGAGATTATTGGCAGAAACATGCGAGAAATGCAATGCTCTGTGAAC | 1440 |
| QY | 1441 | GATAGAAGAAATGTTGGGATCATGATGCAAGACCAAGTTATCAGTGGATGTTAGTGA | 1500 |
| DB | 1441 | GATAGAAGAAATGTTGGGATCATGATGCAAGACCAAGTTATCAGTGGATGTTAGTGA | 1500 |
| QY | 1501 | AGAAATTTACTCAGATGCAAGACTTAACAAATATCATTTACTACAGAGGACATTTGAACAGT | 1560 |
| DB | 1501 | AGAAATTTACTCAGATGCAAGACTTAACAAATATCATTTACTACAGAGGACATTTGAACAGT | 1560 |
| QY | 1561 | GGTCACAAATGGTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGGGCA | 1620 |
| DB | 1561 | GGTCACAAATGGTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGGGCA | 1620 |
| QY | 1621 | CCCTCTGTACACTGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAAGTGC | 1680 |
| DB | 1621 | CCCTCTGTACACTGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAAGTGC | 1680 |
| QY | 1681 | TTAGTTGATTTTCTGTGTGAAATCAGTAGGATGCTCCAGGACATGTACGCAAGCAGCCC | 1740 |
| DB | 1681 | TTAGTTGATTTTCTGTGTGAAATCAGTAGGATGCTCCAGGACATGTACGCAAGCAGCCC | 1740 |
| QY | 1741 | CTTGTGAAAGCATGGATCTGGGAGATGGATCTGGGAACTTACTGCAATCGTCTGCAGCA | 1800 |
| DB | 1741 | CTTGTGAAAGCATGGATCTGGGAGATGGATCTGGGAACTTACTGCAATCGTCTGCAGCA | 1800 |
| QY | 1801 | CAGATATGAAGAGGAGTCTAAGGAAAGAGCTGCAAACTGTAAAGAACTTCTGAAATGTA | 1860 |
| DB | 1801 | CAGATATGAAGAGGAGTCTAAGGAAAGAGCTGCAAACTGTAAAGAACTTCTGAAATGTA | 1860 |
| QY | 1861 | CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTT | 1920 |
| DB | 1861 | CTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTT | 1920 |
| QY | 1921 | GCAAACTGACATCAGATTTCTTAATGCTCTGTCAGAGACACTAATCTCTTAATGAACCT | 1980 |
| DB | 1921 | GCAAACTGACATCAGATTTCTTAATGCTCTGTCAGAGACACTAATCTCTTAATGAACCT | 1980 |
| QY | 1981 | ACTGCTATTTTTTTTAAATGAAAACTTTTCAATTTTCAATTTTAAAGGGTAACTTTTT | 2040 |
| DB | 1981 | ACTGCTATTTTTTTTAAATGAAAACTTTTCAATTTTCAATTTTAAAGGGTAACTTTTT | 2040 |
| QY | 2041 | ATTGCAATTTGCTGTGTTTCTTATAAATGACTATTGTAATGCCAACATGACACAGCTTGTG | 2100 |

2041 ATTGCATTGCTGTTGTTCTTATATAATGACTATTGTATGCAACATGACACAGCTGTG 2100
2101 AATGTGTAGTGTGCTGCTGTTCTGTGTATCATCAAGTGGGTACAGTAAGAGG 2160
2101 AATGTGTAGTGTGCTGCTGTTCTGTGTATCATCAAGTGGGTACAGTAAGAGG 2160
2161 CTTCCAAGCAATTAATTAACTCCCTCAACAGGTATACCTCAGTTCACCGTGTGTAA 2220
2161 CTTCCAAGCAATTAATTAACTCCCTCAACAGGTATACCTCAGTTCACCGTGTGTAA 2220
2221 TTATAAATGAAAACACTAACAGATTTGAATAATCAGTCCATGTTTATACAGGT 2280
2221 TTATAAATGAAAACACTAACAGATTTGAATAATCAGTCCATGTTTATACAGGT 2280
2281 TAATTACAAATCTAGTGTGTTATTTAAGAAAAATGCTAAGCTATGCTAGTCCCAATAG 2340
2281 TAATTACAAATCTAGTGTGTTATTTAAGAAAAATGCTAAGCTATGCTAGTCCCAATAG 2340
2341 TAAGTGCTATTTGTAAGCAGTGTGTTTCTTCTACTGCTGTTGTAATTTAGGGA 2400
2341 TAAGTGCTATTTGTAAGCAGTGTGTTTCTTCTACTGCTGTTGTAATTTAGGGA 2400
2401 AAACAAGTGTCTTTGTAAGTGGAAAGAAATATGCTGTCACCTACCCCTCATCTAT 2460
2401 AAACAAGTGTCTTTGTAAGTGGAAAGAAATATGCTGTCACCTACCCCTCATCTAT 2460
2461 ATCAAGTCCCAAAATATCTTTTCCATTTCAAAGCAGCCTTTGAAACCCCTAAATTA 2520
2461 ATCAAGTCCCAAAATATCTTTTCCATTTCAAAGCAGCCTTTGAAACCCCTAAATTA 2520
2521 CAAGCCAGTAGAAGAAAAGCTAAACACGCTTTTACAAATAGCC 2563
2521 CAAGCCAGTAGAAGAAAAGCTAAACACGCTTTTACAAATAGCC 2563

RESULT 9
ADB53787
D ADB53787 standard; DNA; 2277 BP.
C ADB53787;
C ADB53787;
T 04-DEC-2003 (first entry)
C Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4329.
C toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
C toxicity marker; toxicity progression; drug screening;
C primary rat hepatocyte toxicity modelling; gene; ds.
C Rattus norvegicus.
C WO2003065993-A2.
C 14-AUG-2003.
C 04-FEB-2003; 2003WO-US003482.
C 04-FEB-2002; 2002US-0353171P.
C 13-MAR-2002; 2002US-0363534P.
C 08-APR-2002; 2002US-0370248P.
C 10-APR-2002; 2002US-0371134P.
C 10-APR-2002; 2002US-0371135P.
C 10-APR-2002; 2002US-0371150P.
C 11-APR-2002; 2002US-0371413P.
C 19-APR-2002; 2002US-0371601P.
C 19-APR-2002; 2002US-0373602P.
C 22-APR-2002; 2002US-0374139P.
C 08-MAY-2002; 2002US-0378370P.
C 09-MAY-2002; 2002US-0378652P.
C 09-MAY-2002; 2002US-0378653P.
C 09-MAY-2002; 2002US-0378665P.
C 09-JUL-2002; 2002US-0394230P.

PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX (GENE-) GENE LOGIC INC.
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
XX Elashoff M;
XX WPI; 2003-731472/69.
XX Determining if a compound induces a toxic effect on a tissue or cell, for
XX identifying hepatotoxic compounds, comprises comparing a gene expression
XX profile of a tissue or cell sample to a database of Tox mean and non-Tox
XX mean values.
XX Claim 44; SEQ ID NO 4329; 874pp; English.
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity
XX markers for the prediction or identification of the physiological state
XX of tissue or cell sample that has been exposed to a compound, or to
XX identify or predict the toxic effects of a compound or an agent. These
XX may also be used as markers for monitoring toxicity progression or for
XX drug screening. The present sequence represents a primary rat hepatocyte
XX toxicity modelling related gene sequence from the present invention.
XX SQ Sequence 2277 BP; 671 A; 428 C; 537 G; 640 T; 0 U; 1 Other;
Query Match 81.2%; Score 2080.4; DB 9; Length 2277;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 2185; Conservative 2; Mismatches 80; Indels 19; Gaps 4;
QY 57 GCGCCTCGGAAATGGAGCTGCTGCAAGTGGCGTTCGCGCTTCTTCTCTCT 116
DB 8 GCGCCTCGGAAATGGAGCTGCTGCAAGTGGCGTTCGCGCTTCTTCTCTCT 67
QY 117 GCTCTTCAGTGTCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTA 176
DB 68 GCTCTTCAGTGTCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTA 127
QY 177 ATTGGAAAAGACAGAACCAACACAGACTGCTGTGAACTTGTATGGTGATAAGATA 236
DB 128 ATTGGAAAAGACAGAACCAACACAGACTGCTGTGAACTTGTATGGTGATAAGATA 187
QY 237 AACGCGCAGATTGTTTGTCTACCTGGAAGATATTTCTGGTTCATTTGAATAGTGAAGC 296
DB 188 AACGCGCAGATTGTTTGTCTACCTGGAAGATATTTCTGGTTCATTTGAATAGTGAAGC 247
QY 297 AAGTTGTTGGCTGGATGATATCACTGTATGACAGGACTGATGTTGTCGAAAGAAAG 356
DB 248 AAGTTGTTGGCTGGATGATATCACTGTATGACAGGACTGATGTTGTCGAAAGAAAG 307
QY 357 ACAGCCCTGAAGTGTACTTTTGTGCTGTGAGGCAATATGTTAATGAAAGTCTCTT 416
DB 308 ACAGCCCTGAAGTGTACTTTTGTGCTGTGAGGCAATATGTTAATGAAAGTCTCTT 367
QY 417 ATTTTCCGGAGATGGAAGTACACAGCCCACTTTCAATCTGTTTACCCGAGCCACT 476
DB 368 ATTTTCCGGAGATGGAAGTACACAGCCCACTTTCAATCTGTTTACCCGAGCCACT 427
QY 477 ATTACAAATCTGCTGTATTCCTGTTATCCCTGTTATGTTAATGTCAGGAATGTCATT 536
DB 428 ACTACAACTCTGCTGTATTCCTGTTATCCCTGTTATGTTAATGTCAGGAATGTCATT 487
QY 537 GTGCAATTTGGGTGTACAGACATCAAGATGCCTACCTCCCTGTTACTGTTCTCTCTC 596

| | | | |
|----|------|--|------|
| Ds | 488 | GTGCGTTTGGGTGTACAGACATCAAGAATGGGCTACCCCTCTGTACTTGTTCCTACTC | 547 |
| Qy | 597 | AAGACCCAGACACCCACCTTCCCATTACTAGGTTGAAGCCATTGACGCTGTTAG | 656 |
| Ds | 548 | AGGACCTTGGACCACTCCACCTTCCCATTACTAGGTTGAAGCCATTGACGCTATTGG | 607 |
| Qy | 657 | AAGTGAAGCAAGGGGAAGATTGGTTGTCTGGAAAGCCAGTTGCTCAATGAATATG | 716 |
| Ds | 608 | AAGTGAAGCAAGGGGAAGATTGGTTGTCTGGAAAGCCAGTTGCTCAATGAATATG | 667 |
| Qy | 717 | TGCTCTCAAAATATTTCCTAATACAGGACCAACAGTCTCTGCGAATGAATATGAAGTCT | 776 |
| Ds | 668 | TGCTCTCAAAATATTTCCTAATACAGGACCAACAGTCTCTGCGAATGAATATGAAGTCT | 727 |
| Qy | 777 | ATAGTCTACCTCGAATGAAGCATGAGAACATACATACAGTTTCATTTGGTGCAGAGAAAGAG | 836 |
| Ds | 728 | ATAGTCTACCTCGAATGAAGCATGAGAACATACATACAGTTTCATTTGGTGCAGAGAAAGAG | 787 |
| Qy | 837 | GCACAGTGTGGATGTGGACCTGTGGCTAATACAGACATTTTCAGAAAGGCTCACTGT | 896 |
| Ds | 788 | GCACAGTGTGGATGTGGACCTGTGGCTAATACAGACATTTTCAGAAAGGCTCACTGT | 847 |
| Qy | 897 | CAGACTTCTTAAGGCTAATGTGGTCTCTTGGAAATCACTTTGTGCATATTCAGAAACCA | 956 |
| Ds | 848 | CAGACTTCTTAAGGCTAATGTGGTCTCTTGGAAATCACTTTGTGCATATTCAGAAACCA | 907 |
| Qy | 957 | TGCTAGAGGATTGGCATATTTACATGAGGATATACCTGGCTTAAAGATGSCCAACAGC | 1016 |
| Ds | 908 | TGCTAGAGGATTGGCATATTTACATGAGGATATACCTGGCTTAAAGATGSCCAACAGC | 967 |
| Qy | 1017 | CTGCAATCTCTCAGGGACATCAAAAGTAAAGATGTGCTGTGAAACCAATCTGACAG | 1076 |
| Ds | 968 | CTGCAATCTCTCAGGGACATCAAAAGTAAAGATGTGCTGTGAAACCAATCTGACAG | 1027 |
| Qy | 1077 | CTTGCAATCTCTGAGTTGGGTTGGCTTAAAGTTCAGGCTGGCAAGTCTGACGTGACA | 1136 |
| Ds | 1028 | CTTGCAATCTCTGAGTTGGGTTGGCTTAAAGTTCAGGCTGGCAAGTCTGACGTGACA | 1087 |
| Qy | 1137 | CCATGCGCAGGTTGTACCCGAGGATATAGGCTCCAGAGGTTGTGAGGGTGTATATA | 1196 |
| Ds | 1088 | CCATGCGCAGGTTGTACCCGAGGATATAGGCTCCAGAGGTTGTGAGGGTGTATATA | 1147 |
| Qy | 1197 | ACTTCCAAAGGGACGCAATTTCTGAGGATAGATATGTACGCCATGGGATTAGTCCTATGG | 1256 |
| Ds | 1148 | ACTTCCAAAGGGACGCAATTTCTGAGGATAGATATGTACGCCATGGGATTAGTCCTATGG | 1207 |
| Qy | 1257 | AATGGCTTCTCGTGCACCTGCTGCAGATGGACCCGTAGATGAGTACATGTTTACCATTGG | 1316 |
| Ds | 1208 | AATGGCTTCTCGTGCACCTGCTGCAGATGGACCCGTAGATGAGTACATGTTTACCATTGG | 1267 |
| Qy | 1317 | AGGAAGAAATGGCAGCATCCATCTCTTGAAGATATGACGGAAGTTGTTGTGCATATAA | 1376 |
| Ds | 1268 | AGGAAGAAATGGCAGCATCCATCTCTTGAAGATATGACGGAAGTTGTTGTGCATATAA | 1327 |
| Qy | 1377 | AAAAGAGGCTGTTTAAAGATTTATGGCAGAACATGCGAGGATGCAATGCTCTGTG | 1436 |
| Ds | 1328 | AAAAGAGGCTGTTTAAAGATTTATGGCAGAACATGCGAGGATGCAATGCTCTGTG | 1387 |
| Qy | 1437 | AAACGATAGAAGATGTTGGGATCATGATGAGAACCCAGGTTATCAGCTGGATGTGTAG | 1496 |
| Ds | 1388 | AAACGATAGAAGATGTTGGGATCATGATGAGAACCCAGGTTATCAGCTGGATGTGTAG | 1447 |
| Qy | 1497 | GTGAAGAAATTTACTCAGATGCAAGACGTTAAACAAATATCATTTACTACAGAGGACATTGTAA | 1556 |
| Ds | 1448 | GTGAAGAAATTTACTCAGATGCAAGACGTTAAACAAATATCATTTACTACAGAGGACATTGTAA | 1507 |
| Qy | 1557 | CAGTGTCAAAATGGTGAAGATTTGACCTTCTCCAAAGAAATCTAGTCTATGATGTT | 1616 |
| Ds | 1508 | CAGTGTCAAAATGGTGAAGATTTGACCTTCTCCAAAGAAATCTAGTCTATGATGTT | 1567 |
| Qy | 1617 | GCACCGCTCTGACACTGAGGCTGGGACTCTGAACTGGAGCTCTAAAGCTAAGGAAA | 1676 |
| Ds | 1568 | TGCACCATCTGTCCACACTGAGAAATCGGGACTCTGAACTGGAGCTCTAAAGCTAAGGAAA | 1627 |

| | | | |
|----|------|--|------|
| Qy | 1677 | GTGCTTAGTTGATTTTCTGTGTAATGAGTAGGATGCTCCAGGACATGTACGCAAGCA | 1736 |
| Ds | 1628 | CTGCTTAGTTTATTTTCTGTGTAATGAGTAGGATGCTCCGCGACATGTATGCAAGCA | 1687 |
| Qy | 1737 | GCCCTTGTGGAAGCATGATCTGGGAGATGGATCTGGGAAAATTACTGCAATCGTCTGC | 1796 |
| Ds | 1688 | GCCCTTGTGGAAGCATGAT-----TGGGAGACTTCTCTGCAAGNCTCTGC | 1733 |
| Qy | 1797 | AGCACAGATATGAGGAGTCTTAAGGGAAAAGCTGCAAACTGTAAAGAACTTCTGAAA | 1856 |
| Ds | 1734 | AACAGGATATGAGGGGGTCTTAAGGGAAA--CTGCGAACTGTAAAGAACTTCTGAAA | 1791 |
| Qy | 1857 | TGTACTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAATCAAGT | 1916 |
| Ds | 1792 | CTTACAGGAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAATCAAGT | 1851 |
| Qy | 1917 | ATTGCAAACTGACATCAGATTTCTTAAATGTCTGTGCAAGACACTAATTCCTTAAATG | 1976 |
| Ds | 1852 | ATTG-AAAACACTGACATCAGATTTCTTAAATGTCTGTGCAAGACACTAATTCCTTAAATG | 1910 |
| Qy | 1977 | AACACTGCTATTTTAAATGAAAACCTTTTCAATTCAGATTTTAAAGGGTAACT | 2036 |
| Ds | 1911 | AACACTGCTATTTTAAATGAAAACCTTTTCAATTCAGATTTTAAAGGGTAACT | 1970 |
| Qy | 2037 | TTTATTTGCAATTTGCTGTTTCTATAAATGACTATTGTAATGCAACATGACACAGCT | 2096 |
| Ds | 1971 | TTTATTTGCAATTTGCTGTTTCTATAAATGACTATTGTAATGCAACATGACACAGCT | 2030 |
| Qy | 2097 | TGTGAATGTGTAGTGTCTGTTTCTGTGTACATAGTCATCAAAAGTGGGGTACAGTAA | 2156 |
| Ds | 2031 | TGTGAATGTGTAGTGTCTGTTTCTGTGTACATAGTCATCAAAAGTGGGGTACAGTAA | 2090 |
| Qy | 2157 | GAGCTTCCAGCATTTACTTTAACTCCCTCAAGGTATACCTCAGTTCCACGGTTGC | 2216 |
| Ds | 2091 | GAGCTTCCAGCATTTACTTTAACTCCCTCAAGGTATACCTCAGTTCCACGGTTGC | 2150 |
| Qy | 2217 | TAAATTATAAAATTTGAAAACACTAACAGAAATTTGAAATAAATCAGTCCATGTTTATAACA | 2276 |
| Ds | 2151 | TAAATTATAAAATTTGAAAACACTAACAGAAATTTGAAATAAATCAGTCCATGTTTATAACA | 2210 |
| Qy | 2277 | AGCTTAATACAAATTCACCTGTGTTTATTAAGAAAAAATGTAAGCTATGCTTAGTGCCA | 2336 |
| Ds | 2211 | AGCT--ATTACAACTCAGTGTGTTTATTAAGAAAAAATGTAAGCTATGCTTAGTGCCA | 2210 |
| Qy | 2337 | ATAGTA 2342 | |
| Ds | 2269 | ATAGAA 2274 | |

RESULT 10
AAT84517
ID AAT84517 standard; cDNA; 2122 BP.
XX
AC AAT84517;
XX
XX 02-DEC-1997 (first entry)
XX
DE Mouse activin receptor cDNA.
XX
KW Activin receptor; neuronal cell-specific receptor; mouse;
neurodegenerative disease; diagnosis; gene therapy; ds.
XX
OS Mus musculus.
PH Key Location/Qualifiers
FT CDS 9..1574
FT /tag= a
FT /note= "Claim 9, the ATG initiation codon may be replaced
FT by hydrogen"
FT misc_feature 537..560
FT /*tag= b
FT /note= "24 bp insert at activity receptor type IIA gene

694 CCAGGAGCACCACCTTCTCCATTAAGTCTGAGTGGTGGAAACCACTGCAATTAAGAGT 753
661 GAAAGCAAGGGGAAAGATTGGTGTGTCTGGAAGCCCAAGTTGCTCAATGAATATGTGGC 720
754 GAAAGCAAGGGGAAAGATTGGTGTGTCTGGAAGCCCAAGTTGCTCAATGAATATGTGGC 813
721 TGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGGAGATGAATATGAAGTCTATAG 780
814 TGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGGCAAAATGAATACGAAGTCTACAG 873
781 TCTACCTGGAAATGAAGCATGAGAACATCTACAGTTCATTGGTGCAGAGAAAGAGGCAC 840
874 TTTGCCCTGGAAATGAAGCATGAGAACATCTACAGTTCATTGGTGCAGAGAAAGAGGCAC 933
841 CAGTGTGGATGTGACCTGTGGCTAATACAGCATTTTCATGAAGAGGCTCACTGTGAGA 900
934 CAGTGTGGATGTGACCTGTGGCTAATACAGCATTTTCATGAAGAGGCTCACTATACAGA 993
901 CTTTCTTAAGCCTAATGTGGTCTCTTGGAAATCACTTCTCATATTCAGAGAAACCAATGGC 960
994 CTTTCTTAAGCCTAATGTGGTCTCTTGGAAATCACTTCTCATATTCAGAGAAACCAATGGC 1053
961 TAGAGGATTTGCAATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAAAGCTTGC 1020
1054 TAGAGGATTTGCAATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAAACCTGC 1113
1021 AATCTCTCACAGGACATCAAAAGTAAATGTCTGTTGAAACCAATCTGACAGCTTG 1080
1114 CATATCTCACAGGACATCAAAAGTAAATGTCTGTTGAAACCAATCTGACAGCTTG 1173
1081 CATTCGCTGACCTTTGGGTTGGCTTAAATTTGAGGCTGCGAATCTGCGAGGCAATCCCA 1140
1174 CATTCGCTGACCTTTGGGTTGGCTTAAATTTGAGGCTGCGAATCTGCGAGGCAATCCCA 1233
1141 TGGCAGGTTGGTATCCCGAGGATATATGCTCCAGAGGTTGGAGGGTGTCTATAAATCTT 1200
1234 TGGCAGGTTGGTATCCCGAGGATATATGCTCCAGAGGTTGGAGGGTGTCTATAAATCTT 1293
1201 CCAAGGAGACGATTTCTGAGGATAGATATGATGCGCATGGGATTAAGTCTGAGGAAAT 1260
1294 CCAAGGAGATGCATTTTGGAGATAGATATGATGCGCATGGGATTAAGTCTGAGGAAAT 1353
1261 GGCCTCTGCTGATCTGCTGAGATGAGGACCGCTGATGATGATGATGATGATGATGATG 1320
1354 GGCCTCTGCTGATCTGCTGAGATGAGGACCGCTGATGATGATGATGATGATGATGATG 1413
1321 AGAAATTTGGCCAGCATCCATCTCTTTGAAGATATGAGGAGGTTGTTGTCATATAAAAAA 1380
1414 GGAATTTGGCCAGCATCCATCTCTTTGAAGATATGAGGAGGTTGTTGTCATATAAAAAA 1473
1381 GAGGCTGTTTAAAGATATTTGGCAGAAACATGCAAGATGGCAATGCTCTGTGAAC 1440
1474 GAGGCTGTTTAAAGATATTTGGCAGAAACATGCTGGAATGGCAATGCTCTGTGAAC 1533
1441 GATAGAGAAATTTGGGATCATGATGCAAGCCAGGTTATCAGCTGGATGTGTAGTGA 1500
1534 CATTTGAAGAAATTTGGGATCATGATGCAAGCCAGGTTATCAGCTGGATGTGTAGTGA 1593
1501 AAGAAATTTACTGAGATGCAAGATCAACAAATCATTTACTACAGAGACATTTGTAACAGT 1560
1594 AAGAAATTTACTGAGATGCAAGATCAACAAATCATTTACTACAGAGACATTTGTAACAGT 1653
1561 GGTCAATGTTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGGTGCA 1620
1654 GGTCAATGTTGACAAATGTTGACTTTCTCCCAAGAAATCTAGTCTATGATGGTGCG 1713
1621 CCGTCTGTACACATGTGAGACTGGGACTCTGAATCTGAGCTGCTAAGCTAAGGAAATGTC 1680
1714 CCACTGTGCACTAAGAAATGGGACTCTGAATCTGAGCTGCTAAGCTAAGGAAATGTC 1773
1681 TT--AGTTGATTTTCTGTGTAATAGATGAGTATGCTCCAGACATGTACGCAAGCAGC 1738
1774 TTACAGTTTATTTTCTGTGTAATAGATGAGTATGCTCTCTTGGAAATGTTTAAAGAAAG 1833

QY 1739 CCTTTGTGGAAGCATGGATCTGGAGATGGATCTGGGAAACTTACTGCAATGCTGTCAG 1798
Db 1834 ACCCTTTGTTGAAA-----ATGTTGCTCTGGGAGACTTACTGCAATGCTGTCAG 1883
QY 1799 CACAGATGATGAG---AGAGTCTAAGGGAAGAGCTGCAAACTGTA-----AAGAACTTC 1850
Db 1884 CACAGATGATGAGGACATGAGACTAAGAGAAACCTTTGCAAACTCTATAAGAAACTTTTG 1943
QY 1851 TGAATAATGCTACTCGAAGAAATGTGGCCCTCCAAATCAAGGATCTTTTGACCTGGCTAA 1910
Db 1944 AAAAGTGTACATGAAGAAATGTAGCCCTCCAAATCAAGGATCTTTTGACCTGGCTAA 2003
QY 1911 TCAAGTATTTGCAAACTGACATCAGATTTCTTAATGTCTGTGAGAGACACATTAATTCCT 1970
Db 2004 TGGAGTGTGTTG-AAAATGACATCAGATTTCTTAATGTCTGTGAGAGACACATTAATTCCT 2062
QY 1971 TAAATGAATCTACTGCTATTTTAAATGAAACCTTTTCAATTCAGATTTTAAAAAGG 2030
Db 2063 TAAATGAATCTACTGCTATTTTAAATGAAACCTTTTCAATTCAGATTTTAAAAAGG 2122
QY 2031 GTAAAC---TTTTTATTTGCAATTTGCTGTTCTATATAATGACTATTTGTAATGCCAAT 2087
Db 2123 GTAACTTGTGTTTATTTGCAATTTGCTGTTCTATATAATGACTATTTGTAATGCCAAT 2182
QY 2088 GACACAGCTTGTGAATGTGTAGTGTCTGTCTGTCTGTGTACAT--AGTCATCAAAGTGG 2145
Db 2183 GACACAGCTTGTGAATGTGTAGTGTCTGTCTGTGTGTACATTAAGTCAATCAAAGTGG 2242
QY 2146 GGTACAGTAAAGAGGCTTCCAAGCATTTACTTTAACTCCCTCAACAGGATATACCTCAGT 2205
Db 2243 GGTACAGTAAAGAGGCTTCCAAGCATTTACTTTAACTCCCTCAACAGGATATACCTCAGT 2302
QY 2206 TCCAGGTTGTTAAATATATAAATTTGAAACACTAACAGATTTGAAATTAATCAAGTCCAT 2265
Db 2303 TCCAGGTTGTTAAATATATAAATTTGAAACACTAACAGATTTGAAATTTGATA-----AT 2353
QY 2266 GTTTTATAACAAGGTTAATTACAAATTTCACTGTTGTTATTAA--GAAAAATGTAAGCTA 2324
Db 2354 AATTCGATCCATGTTTGTACAAATTTCACTGTTGTTATTAAAGAAAAAAGGTAAGCTA 2413
QY 2325 TGCCTAGTGGCAATAGTAAAGTGGCTATTTGTAAGCAGTGTGTTTGTCTTCTTCTACTG 2384
Db 2414 TGCCTAGTGGCAATAGTAAAGTGGCTATTTGTAAGCAGTGTGTTTGTCTTCTTCTACTG 2473
QY 2385 GCTTGTAAATTTAGGAAAAACAAGTGTCTGCTTTTGAATAAGGA 2425
Db 2474 GCTTGTAAATTTAGGAAAAACAAGTGTCTGCTTTTGAATAAGGA 2514

RESULT 12
AAS94867
ID AAS94867 standard; DNA; 5286 BP.
XX AAS94867;
AC AAS94867;
XX 14-FEB-2002 (first entry)
DT 14-FEB-2002 (first entry)
XX Human DNA sequence #122 expressed during foam cell differentiation.
DE Human; foam cell differentiation; atherosclerosis; cerebral stroke;
XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX Homo sapiens.
OS Homo sapiens.
XX WO200177389-A2.
PN 18-OCT-2001.
XX 18-OCT-2001.
PD 04-APR-2001; 2001WO-US011128.
XX 04-APR-2001; 2001WO-US011128.
PE 05-APR-2000; 2000US-0195106P.
XX 05-APR-2000; 2000US-0195106P.
XX

| | | |
|----|---|--|
| PA | (INCY-) INCYTE GENOMICS INC. | |
| XX | Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GV, Mikita T; | |
| PI | Tai J; | |
| PI | WPI; 2002-010925/01. | |
| XX | Composition useful for diagnosis of conditions, disorders or diseases | |
| PT | associated with atherosclerosis, comprises several polynucleotides that | |
| PT | are differentially expressed in foam cell development. | |
| XX | Claim 1; Page 177-179; 315pp; English. | |
| PS | The present invention relates to the isolation of human polynucleotide | |
| XX | sequences that are differentially expressed during foam cell | |
| CC | differentiation. The polynucleotide sequences of the invention or a | |
| CC | composition comprising these polynucleotides are useful as a high | |
| CC | throughput method for detecting altered expression of one or more | |
| CC | polynucleotides in a sample. The polynucleotides can be used in the | |
| CC | diagnosis of disorders associated with foam cell development such as | |
| CC | atherosclerosis, cerebral stroke, and cardiovascular disorders such as | |
| CC | coronary artery disease. The polynucleotide sequences can also be used as | |
| CC | PCR primers and probes. The polynucleotides of the invention are also | |
| CC | useful in gene therapy. AAS94746-AAS95021 represent the human | |
| CC | polynucleotide sequences of the invention which are differentially | |
| CC | expressed during foam cell differentiation | |
| XX | Sequence 5286 BP; 1608 A; 915 C; 1017 G; 1712 T; 0 U; 34 Other; | |
| SQ | Query Match 77.2%; Score 1978.2; DB 6; Length 5286; | |
| PI | Best Local Similarity 91.0%; Pred. No. 0; | |
| XX | Matches 222; Conservative 2; Mismatches 181; Indels 37; Gaps 10; | |
| QY | 1 CTCGAGGAGACCCAGGAACTGGATATCTAGCGAGAACTTCCTACGGCTTCCTCGGCG 60 | |
| DB | 103 CTCGAGGAGACCCAGGAACTGGATATCTAGCGAGAACTTCCTCGGANTCCCCGGCG 162 | |
| QY | 61 CCTCGGAAAAATGGAGCTGCTGCAAAAGTTGGCGTTGGCGCTTTCTTTATCTCTTGCTC 120 | |
| DB | 163 CCTCGGAAAAATGGAGCTGCTGCAAAAGTTGGCGTTGGCGCTTTCTTTATCTCTTGCTC 222 | |
| QY | 121 TTCAGGTGCTATACATTGGCAGATACAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTG 180 | |
| DB | 223 TTCAGGTGCTATACATTGGTGTAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTG 282 | |
| QY | 181 GGAARACAGACCAACACCACTGGTGTGAACTGCTATGCTGATGCTGATGATGATGATGATG 240 | |
| DB | 283 GGAARACAGACCAACCACTGGTGTGAACTGCTATGCTGATGCTGATGATGATGATGATGATG 342 | |
| QY | 241 CGGCATTGTTTG-CTACCTGGAGAAATATTTCTGGTTCCATTGAAATAGTGAAGCAAG 299 | |
| DB | 343 CGGCATTGTTTGCTACCTGGAGAAATATTTCTGGTTCCATTGAAATAGTGAAGCAAG 402 | |
| QY | 300 GTTGTGGCTGGATGATATCACTGCTATGACAGGACTGATGTTGTTGTTGTTGTTGTTGTTG 359 | |
| DB | 403 GTTGTGGCTGGATGATATCACTGCTATGACAGGACTGATGTTGTTGTTGTTGTTGTTGTTG 462 | |
| QY | 360 GCCCTGAAGTGTACTTTTCTGCTGAGGGCAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTT 419 | |
| DB | 463 GCCCTGAAGTGTACTTTTCTGCTGAGGGCAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTT 522 | |
| QY | 420 TTCGGAGATGGAAGTACACAGCCCACTTCAATCTGTTTACACCGAAGCCACCTTATT 479 | |
| DB | 523 TTCGGAGATGGAAGTACACAGCCCACTTCAATCTGTTTACACCGAAGCCACCTTATT 582 | |
| QY | 480 ACAACATTCTGCTGATATCTCTTGGTACCACTAATGTTAATTGTCAGGAATTTGTCATTGTTG 539 | |
| DB | 583 ACAACATCTGCTGCTATCTCTTGGTGCCTTATGTTAATTGTCAGGAATTTGTCATTGTTG 642 | |
| QY | 540 CATTTTGGGTGTACAGACATACAGAGTGGCCTACCCCTCTGCTACTTGTCTCTACTCAAG 599 | |
| DB | 643 CATTTTGGGTGTACAGGCAATCAAGATGGCCTACCCCTCTGCTACTTGTCTCTACTCAAG 702 | |

Db 1793 CTTACAGTTATTTCTGTGTAATAGTAGGATGCTCTTGGAAATGTTAAGAAAGAA 1842
Qy 1738 CCCCTTGTGGAAGCATGATCGGAGATGAGTCTGGGAAACTTACTGATCGTCTGCA 1797
Db 1843 GACCCCTTGTGTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1892
Qy 1798 GCACAGATGATGAG--AGGAGTCTAAGGGAAGAGTCTGCAAACTGTA-----AAGAACTT 1849
Db 1953 GCACAGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1952
Qy 1850 CTGAAATGATCTGCAAGAGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1909
Db 1953 GAAAGAGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2012
Qy 1910 ATCAAGTATTTCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1969
Db 2013 ATGAGTGTGTTG-AAAAGTATGATGATGATGATGATGATGATGATGATGATGAT 2071
Qy 1970 TTAATGAACTACTGCTATTTTAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 2029
Db 2072 TTAATGAACTACTGCTATTTTAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 2131
Qy 2030 GGTAACT--TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2086
Db 2132 GGTAACTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2191
Qy 2087 TGACACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2144
Db 2192 TGACACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2251
Qy 2145 GGTACAGTAAAGAGGCTTCAAGCATTAATTAAGTATTAAGTATTAAGTATTAAGTAT 2204
Db 2252 GGTACAGTAAAGAGGCTTCAAGCATTAATTAAGTATTAAGTATTAAGTATTAAGTAT 2311
Qy 2205 TTCCAGGCTGTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2264
Db 2312 TTCCAGGCTGTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2362
Qy 2265 TGTATTATAACAGGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2323
Db 2363 TAAATCGATCCATGTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2422
Qy 2324 ATGCTTAGTGCATAGTAAAGTGGCTATTTGTAAGCAGTGTTTGTAGCTTTTCTTCT 2383
Db 2423 ATGCTTAGTGCATAGTAAAGTGGCTATTTGTAAGCAGTGTTTGTAGCTTTTCTTCT 2482
Qy 2384 GGCCTGTAATTTAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2425
Db 2483 GGCCTGTAATTTAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2524

RESULT 13
AB235028
ID ID
AB235028 standard; cDNA; 2382 BP.
XX
AC AB235028;
XX
DT 05-FEB-2003 (first entry)
XX
DE Human gene expression profile polynucleotide SEQ ID NO 140.
XX
KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW gene expression; gene; ss.
XX
DS Homo sapiens.
XX
PN W0200274979-A2.
XX
PD 26-SEP-2002.
XX

PF 20-MAR-2002; 2002WO-US008456.
PR 20-MAR-2001; 2001US-0276947P.
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
PI Wan J, Wang Y;
XX WPI; 2002-740862/80.
DR
XX
PT New gene expression profile generated from primary, endothelial,
PT epithelial, and muscle cell types, useful for identifying disease
PT pathologies involving alterations of gene expression, e.g. cancer.
XX
PS Claim 3; Page 371-372; 850pp; English.
XX
CC The invention relates to a gene expression profile comprising one or more
CC genes (AB234899-AB235692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,
CC treatment methods, and prognosis. The gene or protein expression profile
CC may also be used for creating microarrays. The microarray is useful for
CC genetic and physical mapping of genomes, DNA sequencing, genetic or
CC medical diagnosis, genotyping of organisms, confirming cell or tissue
CC identifications and in identifying promising antibiotics, antiviral or
CC antifungal agents
XX
SQ Sequence 2382 BP; 698 A; 464 C; 551 G; 669 T; 0 U; 0 Other;
Query Match 74.2%; Score 1903; DB 6; Length 2382;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 2089; Conservative 2; Mismatches 153; Indels 26; Gaps 7;
Qy 1 CTCGAGGAG 60
Db 104 CTCGAGGAG 163
Qy 61 CTCGAGGAG 120
Db 164 CTCGAGGAG 223
Qy 121 TTCAGGTGCTATCTTTCGAGATCAGAACTCAGGAGTGTCTTTTAAATGCTAAATG 180
Db 224 TTCAGGTGCTATCTTTCGAGATCAGAACTCAGGAGTGTCTTTTAAATGCTAAATG 283
Qy 181 GGAAG 240
Db 284 GGAAG 343
Qy 241 GCGCATTTGTTTGTCTTACCTGGAAGATATTTCTGTTTCCATTCAGATAGTGAAGCAGG 300
Db 344 GCGCATTTGTTTGTCTTACCTGGAAGATATTTCTGTTTCCATTCAGATAGTGAAGCAGG 403
Qy 301 TTGTTGGCTGGATGATATCAACTGCTATCAGAGAGTGTGTTGTTGTTGTTGTTGTTGTT 360
Db 404 TTGTTGGCTGGATGATATCAACTGCTATCAGAGAGTGTGTTGTTGTTGTTGTTGTTGTT 463
Qy 361 CCTGAGAGTGTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420

464 CCTGAAGTATATTTTGTGCTGTGAGGGCAATATGTGTAAATGAAAAAGTTTCTTATTT 523
421 TCCGGAGTGAAGTACACAGACCCACTTCAATCCCTGTTTACCCAGAGCCACCCCTATTA 480
524 TCCGGAGTGAAGTACACAGACCCACTTCAATCCCTGTTTACCCAGAGCCACCCCTATTA 583
481 CAACATCTGCTGCTATTTCTTGTGTCACCAATTAATGTAATTCAGGAATGTCATTTGTGC 540
584 CAACATCTGCTGCTATTTCTTGTGTCACCAATTAATGTAATTCAGGAATGTCATTTGTGC 543
541 ATTTTGGGTGTACAGACATACAGATAGGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
644 ATTTTGGGTGTACAGGATACAGATAGGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703
601 CCCAGACACACCCCACTTCCCACTTACTAGGCTTGAAGCCATGACAGCTGTTAGAGT 660
704 CCCAGACACACCCCACTTCTCCATTACTAGGCTTGAAGCCATGACAGCTGTTAGAGT 763
661 GAAAGCAAGGGGAAGTTTGGTGTGTCGTAAGAGCCAGTTGCTCAATGAATATGTGC 720
764 GAAAGCAAGGGGAAGTTTGGTGTGTCGTAAGAGCCAGTTGCTTAAAGCAATATGTGC 823
721 TGTCAAAATATTTTCCAATACAGGACAAACAGTCCCTGCGAGAAATGAATATGAAGTCTATAG 780
824 TGTCAAAATATTTTCCAATACAGGACAAACAGTCCCTGCGAGAAATGAATATGAAGTCTACAG 883
781 TCTACTCTGGAATGAAGCATGAGAAATACATACAGTTCAATGGTGAGAGAAAGAGGCAC 840
884 TTTGCCCTGGAATGAAGCATGAGAAATACATACAGTTCAATGGTGAGAGAAAGAGGCAC 943
841 CAGTGTGATGTGGACCTGTGGCTTAATACAGCATTTTCAATGAAGGCTCACTCTCAGA 900
944 CAGTGTGATGTGGACCTTTTGGCTGATCACAGCATTTTCAATGAAGGCTCACTATCAGA 1003
901 CTTTCTTAAGGCTAAATGTGTCTCTTGGAAATCACTTTGTCTATATGTCAGAAACCATGSC 960
1004 CTTTCTTAAGGCTAAATGTGTCTCTTGGAAATCACTTTGTCTATATGTCAGAAACCATGSC 1063
961 TAGAGATTGGCATATTTACATAGGATATACCTGGCTTAAAGATGCCCACAAAGCTGSC 1020
1064 TAGAGATTGGCATATTTACATAGGATATACCTGGCTTAAAGATGCCCACAAAGCTGSC 1123
1021 AATCTCTCAGGAGCATCAAAAGTAAATATGCTGTGTGTGTAATAATCTGACAGCTTG 1080
1124 CATATCTCAGGAGCATCAAAAGTAAATATGCTGTGTGTGTAATAATCTGACAGCTTG 1183
1081 CATTGCTGCTTTGGTGTGGCTTAAAGTTCGAGGCTGGCAAGTCTGAGGTGACACCCA 1140
1184 CATTGCTGCTTTGGTGTGGCTTAAAGTTCGAGGCTGGCAAGTCTGAGGCGATACCCA 1243
1141 TGGGCAAGTTGGTACCCGGAGGTATATGGCTCCAGAGGTGTGGAGGCTGTATTAACCTT 1200
1244 TGGACAGTTGGTACCCGGAGGTATATGGCTCCAGAGGTATATGGAGGCTGTATTAACCTT 1303
1201 CCAAGGAGCGCATTTCTGAGGATAGATATATGCGCATGGGATTTAGTCTTATGGGAATT 1260
1304 CCAAGGAGCGCATTTTGGAGATAGATATATGATGCCATGGGATTTAGTCTTATGGGAATT 1363
1261 GGCTTCTGTTGCTGCTGCTGAGATGAGCCCGTAGATGATGATGATGATGATGATGATGATG 1320
1364 GGCTTCTGCTGCTGCTGCTGAGATGAGCCCGTAGATGATGATGATGATGATGATGATGATG 1423
1321 AGAAATTTGGCAGCATCCATCTCTTGAAGATATGAGGAGTGTGTGTCATTAATAAAA 1380
1424 GGAAATTTGGCAGCATCCATCTCTTGAAGATATGAGGAGTGTGTGTCATTAATAAAA 1483
1381 GAGGCTGTTTGAAGATTTTGGCAAAATATGCAAAATATGCAAAATATGCAAAATATGCAAA 1440
1484 GAGGCTGTTTGAAGATTTTGGCAAAATATGCAAAATATGCAAAATATGCAAAATATGCAAA 1543
1441 GATAGAGAAATGTTGGGATCATGATGCAAGAGCCAGGTTTATCAGCTGGATGTGTAGGTGA 1500
1544 CATTGAAGATGTTGGGATCAGGACGAGAGCCAGGTTTATCAGCTGGATGTGTAGGTGA 1603

1501 AAGAATTTACTCAGATGCAAGAGCTAAACAATATCATTTACTACAGAGGACATTTGTAACAGT 1560
1604 AAGAATTTACCCAGATGCGAGAGCTAAACAATATTTATTACCAAGAGGACATTTGTAACAGT 1663
1561 GGTCACAATGGTGAACAATGTTGACTTTCTCCCAAGAATCTAGTCTATGATGTTGGCA 1620
1664 GGTCACAATGGTGAACAATGTTGACTTTCTCCCAAGAATCTAGTCTATGATGTTGGC 1723
1621 CCGTCTGTACACTGAGGAGCTGGAGCTGAACTGAGAGCTGCTAAGCTAAGCAAGAGTGC 1680
1724 CCATCTGTGCACACTAAGAAATGGGACTCTGAATGGGAGCTGCTAAGCTAAGCAAGAGTGC 1783
1681 TT--AGTTGATTTTCTGTGTGAATAGTAGAGTCCCTCCAGGACATGTACGCAAGCAGC 1738
1784 TTACAGTTTATTTTCTGTGTGAATAGTAGAGTCTCTTGGAAATGTTAAGAAAGAG 1843
1739 CCCTTGTGGAAGCATGAGTCTGGGAGATGATCTGGGAAATCTTACTGTCATCGTCTGCAG 1798
1844 ACCCTTGTGAAAA-----ATGTTCTCTGGGAGACTTACTGTCATTTGCCGACAG 1893
1799 CACAGATATGAAG--AGGAGTCTAAGGAAAAAGCTGCAAACTGTAAAGAACTTCTTG--- 1852
1894 CACAGATGTGAAGGACATGAGACTAAGAGAAACCTTGCAAACTCTATAAGAAACCTTTTG 1953
1853 --AAATGTACTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAA 1910
1954 AAAAAGTGTACATGAAGATGTAGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAA 2013
1911 TCAAGTATTTGCAAACTGACATCAGATTTCTTAATGTCTGTCAAGACACACTTAATCTCT 1970
2014 TGGAGTGTG-AAAATGACATCAGATTTCTTAATGTCTGTCAAGACACACTTAATCTCT 2072
1971 TAAATGAACACTGCTATTTTAAATGAAAACTTTTCAATTTTCAATTTTAAAGAG 2030
2073 TAATGAACACTGCTATTTTAAATGAAAACTTTTCAATTTTCAATTTTAAAGAG 2132
2031 GTAAC---TTTTATGCAATTTGCTGTTTCTAATAATGACTATTTGTAATGCTCAACAT 2087
2133 GTAATGTTTATTTGCAATTTGCTGTTTCTAATAATGACTATTTGTAATGCTCAACAT 2192
2088 GACACAGCTTGTGATGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2145
2193 GACACAGCTTGTGATGTGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2252
2146 GGTACAGTAAAGAGGCTTCCAGCATTTAACTCCCTCCCAAGGTATACCTCAGT 2205
2253 GGTACAGTAAAGAGGCTTCCAGCATTTAACTCCCTCCCAAGGTATACCTCAGT 2312
2206 TCCACGCTTCTTAAATTTAATAATGAAAACTAACAAGAAATTTGAATAA 2255
2313 TCCACGCTTCTAATTTAATAATGAAAACTAACAAGAAATTTGAATAA 2362

RESULT 14

AAH20184

ID AAH20184 standard; DNA; 1442 BP.

XX

AC AAH20184;

XX

XX 09-AUG-2001 (first entry)

XX

XX Mouse two-hybrid plasmid containing activin IIA DNA SEQ ID NO:1.

XX

XX Mouse; activin receptor affinity; PDZ domain; identification;
KW intracellular activin signal transducing molecule; binding protein;
KW notropin; neuroprotective; antiparkinsonian; anticonvulsant;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW brain disorder; nerve disorder; ds.

XX

OS Mus sp.

XX

XX WO200131004-A1.

PN


```

RESULT 15
AAA59943
ID AAA59943 standard; DNA; 1466 BP.
XX
XX AAA59943;
XX
XX 16-OCT-2000 (first entry)
XX
XX Polynucleotide sequence SEQ ID 1.
XX
XX Activin receptor binding protein; brain; PDZ domain; WW domain; epilepsy; mouse; activin intracellular transducer molecule; Parkinson's disease; nervous system disorder; Alzheimer's disease; Huntington's disease; DS.
XX
XX Mus sp.
XX
XX W0200029570-Al.
XX
XX 25-MAY-2000.
XX
XX 11-NOV-1999; 99WO-JP006275.
XX
XX 13-NOV-1998; 98JP-00323199.
XX
XX 07-DEC-1998; 98JP-00346925.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Sugino H;
XX
XX WPI; 2000-387783/33.
XX
XX Polypeptide with affinity for activin receptor and activin intracellular transducer for treatment of Parkinson's and Alzheimer's diseases and other neurological disorders.
XX
XX Example 1; Page 103-104; 127pp; Japanese.
XX
XX This invention relates to a polypeptide, and its salts, which is highly expressed in brain tissue and contains at least one PDZ and/or WW domain. The polypeptide binds to the activin receptor and/or activin intracellular transducer molecule (AIRM). Also included in the invention is a polynucleotide sequence encoding the polypeptide, and its partial peptides. Vectors containing the nucleotide sequence, host cells transformed by the vector, antibodies recognising the protein, and assays using the antibodies are also included. The activin receptor binding protein is used for the treatment, prevention and diagnosis of nervous system disorders such as Parkinson's disease, Alzheimer's disease, Huntington's disease and epilepsy. This sequence represents a polynucleotide sequence used in the identification of the activin receptor binding protein of the invention
XX
XX Sequence 1466 BP; 442 A; 279 C; 363 G; 382 T; 0 U; 0 Other;
XX
XX Query Match 54.9%; Score 1406; DB 3; Length 1466;
XX
XX Best Local Similarity 98.2%; Pred. No. 0;
XX
XX Matches 1440; Conservative 1; Mismatches 1; Indels 24; Gaps 1
XX
XX
XX
XX 554 AGACATCACAGATGCGCTACCCCTCTGTACTTGTTCCTACTCAA----- 598
XX
XX 1 AGACATCACAGATGCGCTACCCCTCTGTACTTGTTCCTACTCAAACGCCCTTTCATATA 60
XX
XX
XX 599 -----GAGCCAGGACACCCCCACCTTCCCTACTTACTAGGGTTGAAGCCATTGCAG 649
XX
XX 61 ATGATAGAGGACCCAGGACACCCCCACCTTCCCTACTTACTAGGGTTGAAGCCATTGCAG 120
XX
XX
XX 650 CTGTTAGAAGTGAAGCAAGCGGGAAGATTGTTGTCTGGAAGCCAGTTGCTCAAT 709
XX
XX 121 CTGTTAGAAGTGAAGCAAGCGGGAAGATTGTTGTCTGGAAGCCAGTTGCTCAAT 180
XX
XX
XX 710 GAATATGTGGCTGTCAAAATATTTCCATACAGGACAAACAGTCTCTGGCAGATGATAT 769
XX
XX
XX 181 GAAATATGTGGCTGTCAAAATATTTCCATACAGGACAAACAGTCTCTGGCAGATGATAT 240
XX
XX

```

Qy 1850 CTGAAATGTACTCGAAGAATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTA 1909
Db ||||||||||||||||||
Qy 1321 CTGAAATGTACTCGAAGAATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTA 1380
Db ||||||||||||||||||
Qy 1910 ATCAAGTATTTGCAAACTGACATCAGATTCTTAATGTCTGTCAAGACACACTAATTC 1969
Db ||||||||||||||||||
Qy 1381 ATCAAGTATTTGCAAACTGACATCAGATTCTTAATGTCTGTCAAGACACACTAATTC 1440
Db ||||||||||||||||||
Qy 1970 TTAATGAACCTACTGCTATTTTTTTT 1995
Db ||||||||||||||||||
Qy 1441 TTAATGAACCTACTGCTATTTTTTTT 1466
Db ||||||||||||||||||

Search completed: February 26, 2004, 03:03:15
Job time : 940 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 26, 2004, 00:08:26 ; Search time 9674 Seconds
(without alignments)
11483.171 Million cell updates/sec

Title: US-09-742-684A-15
Perfect score: 2563
Sequence: 1 cctcgaggagaccagggga.....aacacgctttacaataagcc 2563

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_hig:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_hig_hum:
31: em_hig_inv:
32: em_hig_other:
33: em_hig_mus:
34: em_hig_pln:
35: em_hig_rtd:
36: em_hig_mam:
37: em_hig_vrt:
38: em_sy:
39: em_higo_hum:
40: em_higo_mus:
41: em_higo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|------------|--------------------|
| 1 | 2556.4 | 99.7 | 2563 | 6 | AR116087 | AR116087 Sequence |
| 2 | 2556.4 | 99.7 | 2563 | 6 | AR136937 | AR136937 Sequence |
| 3 | 2556.4 | 99.7 | 2563 | 10 | MUSACTR | M65287 Mouse activ |
| 4 | 2080.4 | 81.2 | 2277 | 10 | S48190 | S48190 type II act |
| 5 | 2057.4 | 80.3 | 2122 | 6 | AR116081 | AR116081 Sequence |
| 6 | 2057.4 | 80.3 | 2122 | 6 | E15480 | E15480 cDNA encodi |
| 7 | 1978.2 | 77.2 | 2586 | 6 | AX281713 | AX281713 Sequence |
| 8 | 1905.6 | 74.4 | 2363 | 9 | HSACTR | X82381 H.sapiens m |
| 9 | 1903 | 74.2 | 2382 | 9 | HUMACTRIIA | M93415 Human activ |
| 10 | 1892.4 | 73.8 | 2375 | 9 | HUMACTRIIA | D11770 Homo sapien |
| 11 | 1495.4 | 58.3 | 1563 | 6 | AR116086 | AR116086 Sequence |
| 12 | 1466 | 57.2 | 1543 | 10 | RATACTRII | L10639 Rat activin |
| 13 | 1440 | 56.2 | 1442 | 6 | BD015780 | BD015780 Novel pro |
| 14 | 1440 | 56.2 | 1442 | 6 | BD095083 | BD095083 Novel pro |
| 15 | 1436.6 | 56.1 | 1670 | 4 | BTU43208 | X63208 Bos taurus |
| 16 | 1435.2 | 56.0 | 1573 | 9 | HSACTREC | X63128 H.sapiens m |
| 17 | 1428 | 55.7 | 1640 | 4 | BOVACTRII | L21717 Bos taurus |
| 18 | 1406 | 54.9 | 1466 | 6 | E43123 | E43123 Novel prote |
| 19 | 1365.8 | 53.3 | 1542 | 4 | SHPACTIVR | L19442 Ovis aries |
| 20 | 1162.4 | 45.4 | 3112 | 5 | GGU31222 | U31222 Gallus gall |
| 21 | 1108.2 | 43.2 | 1542 | 5 | CHKARTIIA | D31899 Gallus gall |
| 22 | 1023.4 | 39.9 | 222471 | 10 | AL732317 | AL732317 Mouse DNA |
| 23 | 987.6 | 38.5 | 2315 | 5 | S70930 | S70930 activin rec |
| 24 | 983.4 | 38.4 | 2313 | 6 | AR116088 | AR116088 Sequence |
| 25 | 881.4 | 34.4 | 220307 | 2 | AC105509 | AC105509 Rattus no |
| 26 | 881.4 | 34.4 | 223345 | 2 | AC123324 | AC123324 Rattus no |
| 27 | 843.4 | 32.9 | 1512 | 10 | S3751852 | S37521 activin rec |
| 28 | 828.4 | 32.3 | 1497 | 10 | MUSACTR22 | M93399 Mouse activ |
| 29 | 697.2 | 27.2 | 189353 | 9 | AC009480 | AC009480 Homo sapi |
| 30 | 678.2 | 26.5 | 2335 | 6 | AR136938 | AR136938 Sequence |
| 31 | 670 | 26.1 | 2075 | 5 | S49438 | S49438 activin rec |
| 32 | 667 | 26.0 | 1903 | 5 | GGU31223 | U31223 Gallus gall |
| 33 | 665.8 | 26.0 | 1587 | 9 | HSACTIIRE | X77533 H.sapiens m |
| 34 | 663.6 | 25.9 | 2542 | 5 | XELACR | M88594 Xenopus lae |
| 35 | 631 | 24.6 | 1833 | 4 | BTU57707 | US7707 Bos taurus |
| 36 | 627.2 | 24.5 | 1651 | 6 | BD191971 | BD191971 The use o |
| 37 | 618 | 24.1 | 2041 | 10 | RATACTRRE | M87067 R.norvegicu |
| 38 | 603.8 | 23.6 | 4403 | 5 | BC044131 | BC044131 Danio rer |
| 39 | 602.2 | 23.5 | 3964 | 5 | AF069500 | AF069500 Danio rer |
| 40 | 596.4 | 23.3 | 1628 | 6 | BD191973 | BD191973 The use o |
| 41 | 571.6 | 22.3 | 2515 | 5 | AF001406 | AF001406 Carassius |
| 42 | 564.2 | 22.0 | 1708 | 10 | MUSACTR2B | M84120 Mouse activ |
| 43 | 564.2 | 22.0 | 1722 | 6 | BD191970 | BD191970 The use o |
| 44 | 558.6 | 21.8 | 1699 | 6 | BD191972 | BD191972 The use o |
| 45 | 538.2 | 21.0 | 570 | 10 | S96418 | S96418 activin rec |

ALIGNMENTS

RESULT 1
AR116087
LOCUS AR116087 2563 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 12 from patent US 6132988.
ACCESSION AR116087
VERSION AR116087.1 GI:14096409
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2563)
AUTHORS Sugino,H., Nakamura,T. and Shouji,H.
TITLE DNA encoding a neuronal cell-specific receptor protein
JOURNAL Patent: US 6132988-A 12 17-OCT-2000;
FEATURES Location/Qualifiers

Db 2041 ATTGATTTGCTGTTGTTCTAATAATGACTATGTAAATGCCAATGACACAGCTTGTG 2100
Qy 2101 AATGTTAGTGTGCTGTTCTGTTGATCATAGTATCATCAAGTGGGTACAGTAAAGAGG 2160
Db 2101 AATGTTAGTGTGCTGTTCTGTTGATCATAGTATCATCAAGTGGGTACAGTAAAGAGG 2160
Qy 2161 CTTCAGCATTTACTTTTAACTTCCCTCAACAGGTATACCTAGTTCCAGGTTGTTAAA 2220
Db 2161 CTTCAGCATTTACTTTTAACTTCCCTCAACAGGTATACCTAGTTCCAGGTTGTTAAA 2220
Qy 2221 TTATAAAATTTGAAACACATCAACAGATTTGAAATTAATCACTGTCATGTTTATACAGGT 2280
Db 2221 TTATAAAATTTGAAACACATCAACAGATTTGAAATTAATCACTGTCATGTTTATACAGGT 2280
Qy 2281 TAATTACAAATTTCACTGTGTTTAAAGAAAATGTTAGTATGTTAGTGGCAATAG 2340
Db 2281 TAATTACAAATTTCACTGTGTTTAAAGAAAATGTTAGTATGTTAGTGGCAATAG 2340
Qy 2341 TAAGTGGCTATTGTAAGCACTGTTTAAAGCACTGTTTCTTCTACTGCTTGTAAATTAGGGA 2400
Db 2341 TAAGTGGCTATTGTAAGCACTGTTTAAAGCACTGTTTCTTCTACTGCTTGTAAATTAGGGA 2400
Qy 2401 AAACAAGTCTGCTTCTGTTGAAATGGAAGAAATGTTAGTGTGTCACCTACCCCTACTTAT 2460
Db 2401 AAACAAGTCTGCTTCTGTTGAAATGGAAGAAATGTTAGTGTGTCACCTACCCCTACTTAT 2460
Qy 2461 ATCAAGGTCCTCAAAATATTTCTTTCATTTCAATTTCAAGACAGCACTTTGAAACCCCTAAATTA 2520
Db 2461 ATCAAGGTCCTCAAAATATTTCTTTCATTTCAATTTCAAGACAGCACTTTGAAACCCCTAAATTA 2520
Qy 2521 CAAGCAGTAGAAGAAAGCTAAACACACCTTTTACAAATAGCC 2563
Db 2521 CAAGCAGTAGAAGAAAGCTAAACACACCTTTTACAAATAGCC 2563

RESULT 2
ARI36937
LOCUS ARI36937 2563 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6162896.
ACCESSION ARI36937
VERSION ARI36937.1 GI:14478187
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2563)
AUTHORS Mathews,L.W., Vale,W.W. Jr. and Tsuchida,K.
TITLE Recombinant vertebrate activin receptors
JOURNAL Patent: US 6162896-A 1 19-DEC-2000;
FEATURES
source
1..2563
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 99.7%; Score 2556.4; DB 6; Length 2563;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2557; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 1 CTCGAGGAAGACCCAGGAACTGATATCTAGCGAACTTCTACGGCTTCTCGGGCG 60
Db 1 CTCGAGGAAGACCCAGGAACTGATATCTAGCGAACTTCTACGGCTTCTCGGGCG 60
Qy 61 CCTCGGAAATGGGAGCTGCTGCAAGGTTGGCGTTTCGCGCTTCTTATCTCTGCTC 120
Db 61 CCTCGGAAATGGGAGCTGCTGCAAGGTTGGCGTTTCGCGCTTCTTATCTCTGCTC 120
Qy 121 TTCAGGTCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAAATGCTAATTG 180
Db 121 TTCAGGTCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAAATGCTAATTG 180
Qy 181 GGAAGACAGAACCAACAGACTGGTGTGAAACCTTGCATGTTGATGATAAAGCAACG 240

Db 181 GGAAGACAGAACCAACAGACTGGTGTGAAACCTTGCATGATAAAGCAACG 240
Qy 241 GGCACATTTGTTTGTCTACCTGGAAGAAATATTTCTGTTTCCATTGAAATAGTGAAGCAAGG 300
Db 241 GGCACATTTGTTTGTCTACCTGGAAGAAATATTTCTGTTTCCATTGAAATAGTGAAGCAAGG 300
Qy 301 TTGTTGGCTGGAATATCAACTGCTATGACAGCACTGATTGTATAGAAAAAAGACAG 360
Db 301 TTGTTGGCTGGAATATCAACTGCTATGACAGCACTGATTGTATAGAAAAAAGACAG 360
Qy 361 CCTGAAAGTGTACTTTTGTGTTGAGGCAATATGTGTAAATGAAAGTTCTCTTATTT 420
Db 361 CCTGAAAGTGTACTTTTGTGTTGAGGCAATATGTGTAAATGAAAGTTCTCTTATTT 420
Qy 421 TCCGAGATGGAAGTACACAGCCCACTTCAAAATCTGTGTACACCGAAGCCACCTATTA 480
Db 421 TCCGAGATGGAAGTACACAGCCCACTTCAAAATCTGTGTACACCGAAGCCACCTATTA 480
Qy 481 CAACATTTCTGCTGTTATTTCTGTTACCACTTAATGTTAAATTCAGGAATTTGCTTTGTC 540
Db 481 CAACATTTCTGCTGTTATTTCTGTTACCACTTAATGTTAAATTCAGGAATTTGCTTTGTC 540
Qy 541 ATTTTGGGTGTACAGACATCAAGATGCGCTACCTCTCTGTACTTGTCTCTACTCAAGA 600
Db 541 ATTTTGGGTGTACAGACATCAAGATGCGCTACCTCTCTGTACTTGTCTCTACTCAAGA 600
Qy 601 CCCAGACACCCCACTTCCCATTTAGGTTGAAGCAATTCAGCTGTTAGAGT 660
Db 601 CCCAGACACCCCACTTCCCATTTAGGTTGAAGCAATTCAGCTGTTAGAGT 660
Qy 661 GAAAGCAAGGGGAAGATTGTTGTTGCTGTTGAAAGCCCTGCTCAATGAATATGTCG 720
Db 661 GAAAGCAAGGGGAAGATTGTTGTTGCTGTTGAAAGCCCTGCTCAATGAATATGTCG 720
Qy 721 TGTCAAAATATTTTCCAATACAGCAAAACAGTCTCTGCGCAGAAATGAATGAAGCTATAG 780
Db 721 TGTCAAAATATTTTCCAATACAGCAAAACAGTCTCTGCGCAGAAATGAATGAAGCTATAG 780
Qy 781 TCTACTGGAATGAAGCATGAGAAACATATACAGTTTCAATGTTGTCAGAGAAAGAGGCAC 840
Db 781 TCTACTGGAATGAAGCATGAGAAACATATACAGTTTCAATGTTGTCAGAGAAAGAGGCAC 840
Qy 841 CAGTGTGATGTGGACCTGTGCTTAATACAGCAATTTCAAGAAAGGCTCAGTGTGAGA 900
Db 841 CAGTGTGATGTGGACCTGTGCTTAATACAGCAATTTCAAGAAAGGCTCAGTGTGAGA 900
Qy 901 CTTTCTTAAGCTTAATGTTGCTCTTGGAAATCACTGTTTGTATGTCAGAAACCATGGC 960
Db 901 CTTTCTTAAGCTTAATGTTGCTCTTGGAAATCACTGTTTGTATGTCAGAAACCATGGC 960
Qy 961 TAGAGGATTTGCATATTTACATGAGGATATACCTGCTTAAAGAGTGGCCACAGCCTGC 1020
Db 961 TAGAGGATTTGCATATTTACATGAGGATATACCTGCTTAAAGAGTGGCCACAGCCTGC 1020
Qy 1021 AATCTCTCAGAGGACATCAAAAGTAAATGTGTTGTTGAAAAACAATCTGACAGTTG 1080
Db 1021 AATCTCTCAGAGGACATCAAAAGTAAATGTGTTGTTGAAAAACAATCTGACAGTTG 1080
Qy 1081 CATTTGCTGACTTTGGGTTGGCCTTAAAGTTCAGAGGCTGCAAGTCTGCAAGTGCACCCA 1140
Db 1081 CATTTGCTGACTTTGGGTTGGCCTTAAAGTTCAGAGGCTGCAAGTCTGCAAGTGCACCCA 1140
Qy 1141 TGGGCAAGTGTGATCCCGAGGATATGCTCCAGAGGTTGTCAGAGGCTCTATAAATTT 1200
Db 1141 TGGGCAAGTGTGATCCCGAGGATATGCTCCAGAGGTTGTCAGAGGCTCTATAAATTT 1200
Qy 1201 CCAAGGACGCAATTTCTGAGGATAGATATGACGCAATGGGATTTAGTCTCTATGGGAAT 1260
Db 1201 CCAAGGACGCAATTTCTGAGGATAGATATGACGCAATGGGATTTAGTCTCTATGGGAAT 1260
Qy 1261 GGCCTTCTGTTGCACTGCTGCAAGTGGACCCCTAGATGATGATGATGATGATGATGATG 1320

1261 GGCTTCTCGTTGACATGACCGGTAGATGAGTACATGTTACCATTTAGGA 1320
 1321 AGAAATGGCCAGCATCTCTTGAAGATATGACGAGTGTGTGTCATAAAAAA 1380
 1321 AGAAATGGCCAGCATCTCTTGAAGATATGACGAGTGTGTGTCATAAAAAA 1380
 1381 GAGGCTGTTTAAAGATATTGGCAGAAACATGACGAGGATGCAATGCTCTGTGAAC 1440
 1381 GAGGCTGTTTAAAGATATTGGCAGAAACATGACGAGGATGCAATGCTCTGTGAAC 1440
 1441 GATAGAGAAATGTTGGATCATGATGACGAGGATGCAATGCTCTGTGAAC 1500
 1441 GATAGAGAAATGTTGGATCATGATGACGAGGATGCAATGCTCTGTGAAC 1500
 1501 AGAATTAATCAAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGAT 1560
 1501 AGAATTAATCAAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGAT 1560
 1561 GGTCAATGGTCAAAATGTTGACTTCTCCCAAGAACTAGTCTATGATGTTGCA 1620
 1561 GGTCAATGGTCAAAATGTTGACTTCTCCCAAGAACTAGTCTATGATGTTGCA 1620
 1621 CCTCTGTACACATGAGGACTGGGACTCTGAATGAGGACTGCTAAGCTAAGGAAAGTGC 1680
 1621 CCTCTGTACACATGAGGACTGGGACTCTGAATGAGGACTGCTAAGCTAAGGAAAGTGC 1680
 1681 TTAGTGTATTTCTGTGTGAATGAGTGAATGCTCCAGGACATGACGAGGAGGAGG 1740
 1681 TTAGTGTATTTCTGTGTGAATGAGTGAATGCTCCAGGACATGACGAGGAGGAGG 1740
 1741 CTTGTGAAAGCATGATCTGGAGATGATGCTGGGAACTTACTGCAATGCTGTCAGCA 1800
 1741 CTTGTGAAAGCATGATCTGGAGATGATGCTGGGAACTTACTGCAATGCTGTCAGCA 1800
 1801 CAGATATGAAGAGGATCTAAGGAGAAAGTCTGCAATGAGGAACTTCTGAAATGTA 1860
 1801 CAGATATGAAGAGGATCTAAGGAGAAAGTCTGCAATGAGGAACTTCTGAAATGTA 1860
 1861 CTGAGAAATGTCGCTCTCCAAATCAAGGATCTTTGGACCTGGCTAATCAAGTATTT 1920
 1861 CTGAGAAATGTCGCTCTCCAAATCAAGGATCTTTGGACCTGGCTAATCAAGTATTT 1920
 1921 GCAAAATGATCATGATTTCTTAATGCTGTGCAAGGACATTAATCTTAAATGAACT 1980
 1921 GCAAAATGATCATGATTTCTTAATGCTGTGCAAGGACATTAATCTTAAATGAACT 1980
 1981 ACTGCTATTTTAAATGAAAGTCTTCAATTCAGATTTTAAAGGTAACCTTTT 2040
 1981 ACTGCTATTTTAAATGAAAGTCTTCAATTCAGATTTTAAAGGTAACCTTTT 2040
 2041 ATTGCATTTGCTGTTTCTAATGATTAATGATTAATGATTAATGATTAATGAT 2100
 2041 ATTGCATTTGCTGTTTCTAATGATTAATGATTAATGATTAATGATTAATGAT 2100
 2101 AATGTGTAGTGTGCTGTTCTGTGATCATAGTCAATCAAGTGGGGTACAGTAAAGAGG 2160
 2101 AATGTGTAGTGTGCTGTTCTGTGATCATAGTCAATCAAGTGGGGTACAGTAAAGAGG 2160
 2161 CTTCCAGCATTAATTAACCTCCCTCAACAGGATTAATCAAGTGGGGTACAGTAAAGAGG 2220
 2161 CTTCCAGCATTAATTAACCTCCCTCAACAGGATTAATCAAGTGGGGTACAGTAAAGAGG 2220
 2221 TTATAAATTCGAAACATTAACAGATTTGAATTAATCAATGCTGCTTTTATAACAGGT 2280
 2221 TTATAAATTCGAAACATTAACAGATTTGAATTAATCAATGCTGCTTTTATAACAGGT 2280
 2281 TAATTAACAAATCACTGTGTTTAAAGAAATTAAGTAAAGTATGCTAGTCCCAATAG 2340
 2281 TAATTAACAAATCACTGTGTTTAAAGAAATTAAGTAAAGTATGCTAGTCCCAATAG 2340
 2341 TAAGTGGCTATTTGTAAGAGCATGTTTACCTTCTCTACTGCTGCTTGAATTTAGGGA 2400
 2341 TAAGTGGCTATTTGTAAGAGCATGTTTACCTTCTCTACTGCTGCTTGAATTTAGGGA 2400

2401 AAACAGTGTCTGCTTTTGAATGAAAGATATGGTGTACCTACCTACCTACCTAT 2460
 2401 AAACAGTGTCTGCTTTTGAATGAAAGATATGGTGTACCTACCTACCTACCTAT 2460
 2461 ATCAAGGTCCCAAAATATCTTTTCCATTTCAAGAGACGACTTTGAAACCTAAATTA 2520
 2461 ATCAAGGTCCCAAAATATCTTTTCCATTTCAAGAGACGACTTTGAAACCTAAATTA 2520
 2521 CAAGCCAGTAGAAGAAAGCTAAACACAGCTTTTCAAAATAGCC 2563
 2521 CAAGCCAGTAGAAGAAAGCTAAACACAGCTTTTCAAAATAGCC 2563

RESULT 3
 MUSACTR
 LOCUS Mouse activin receptor (Actr) mRNA, complete cds. ROD 27-APR-1993
 DEFINITION M65287
 ACCESSION M65287.1 GI:191663
 VERSION
 KEYWORDS activin receptor; serine kinase.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (Bases 1 to 2563)
 AUTHORS Mathews, L.S. and Vale, W.W.
 TITLE Expression cloning of an activin receptor, a predicted transmembrane serine kinase
 JOURNAL Cell 65 (6), 973-982 (1991)
 MEDLINE 91256317
 PUBMED 1646080
 COMMENT Original source text: Mus musculus cDNA to mRNA.
 FEATURES
 Location/Qualifiers
 1..2563
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /cell_line="Act20"
 1..2563
 /gene="Actr"
 71..1612
 /gene="Actr"
 /codon_start=1
 /product="activin type IIB receptor"
 /protein_id="AAA37171.1"
 /db_xref="GI:191664"
 /translation="MGAAKLAFAVFLISCSGAILGRSETQCLFNFANWERDRNQ
 TGVEPCYGDKDRSHCEATWNIISGSIIVKQGWLDINDCYDRDTCIEKQSPVEYF
 CCCEGNMCNEKFSYFPEMEVTPQSNPTPKPPYNNILLYSLVPLMLIAGIVLCYFW
 YRHKKMAYPPVLVPTQDPQPPPPPLGLKQLLELVKARGFGCVKQALLNEYVAV
 KIPFIQKQWQNEVEYVSLFGMKHENILQFIGAEKRGTSVDVLDLWLTAFHEKGSUS
 DFLKANVSWNELCHIAETWARGLYLHEDIPGLADKHKPAISHREDIKSNLLKNNL
 TACIADFLGALKPEAGSAGDTGQVTRRYMAPEVLGAINFQRFDAFLRIDMYAMGL
 VLWELASRCLTAADGPDVDEMLPFSEIIGHPSLEDMQEVVHKKRPVLRDVMYKHAG
 NMALCEITIECWDHDAEARLSAGCVGERITQMRLTNITTEDIVTVTVTVNVDPPP
 KESSY"
 71..127
 /gene="Actr"
 128..1609
 /gene="Actr"
 /product="activin type IIB receptor"

Query Match 99.7%; Score 2556.4; DB 10; Length 2563;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2557; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CTCGAGGAGAACCCAGGAACTGGATATCTAGCGAGAACTTCTACGGTCTTCGGCG 60
 Db 1 CTCGAGGAGAACCCAGGAACTGGATATCTAGCGAGAACTTCTACGGTCTTCGGCG 60
 QY 61 CTCGAGGAGAACCCAGGAGTGTGCAAGTGGCGTTCGGCGCTTCTTATCTCTGCTC 120

| | | | |
|------|----|--|------|
| 61 | db | CTCTGGGAAAATGGGAGCTGCTGCAAAAGTTGGCGTTGCGCGTCTTCTTATCTCTTGCTC | 120 |
| 121 | 2Y | TTTCAGGTGCTATACTTGGCAGATCAGAACTCAGAGTGTCTTTTCTTTAATGCTAAATG | 180 |
| 121 | db | TTTCAGGTGCTATCTTTGGCAGATCAGAACTCAGAGTGTCTTTTCTTTAATGCTAAATG | 180 |
| 181 | 2Y | GGAAARGACAGAACCAACGACTGGTGTGCAACCTTGCTATGCTGATAAAGATAAAG | 240 |
| 181 | db | GGAAAGACAGAACCAACGACTGGTGTGCAACCTTGCTATGCTGATAAAGATAAAG | 240 |
| 241 | 2Y | GGCAGATTTGTTTGTCTACTCTGGAAGATAATTTCTGGTCCATTTGAATAGTGAAGCAAG | 300 |
| 241 | db | GGCAGATTTGTTTGTCTACTCTGGAAGATAATTTCTGGTCCATTTGAATAGTGAAGCAAG | 300 |
| 301 | 2Y | TTGTTGGCTGGATGATATCAACTGCTATGACAGACTGATTTGTGTGNGAAAAAAGACAG | 360 |
| 301 | db | TTGTTGGCTGGATGATATCAACTGCTATGACAGACTGATTTGTGTGNGAAAAAAGACAG | 360 |
| 361 | 2Y | CCCTGAAGTGTACTTTTGTGCTGTGAGGGCAATATGCTGTAATGAAAGTTCTCTTATTT | 420 |
| 361 | db | CCCTGAAGTGTACTTTTGTGCTGTGAGGGCAATATGCTGTAATGAAAGTTCTCTTATTT | 420 |
| 421 | 2Y | TCCGGAGATGGAAGTCACACGCCACATTCAAATCCTGTTACACGAGGCCACCTTATTA | 480 |
| 421 | db | TCCGGAGATGGAAGTCACACGCCACATTCAAATCCTGTTACACGAGGCCACCTTATTA | 480 |
| 481 | 2Y | CAACATTTCTGCTATTTCTTGGTACCACTAATGTTAAATTGAGGAATTTGTCATTGTGTC | 540 |
| 481 | db | CAACATTTCTGCTATTTCTTGGTACCACTAATGTTAAATTGAGGAATTTGTCATTGTGTC | 540 |
| 541 | 2Y | ATTTTGGTGTACAGACATCACAGAATGGCCCTACCTCTCTACTTTGTTCTTACTCAAGA | 600 |
| 541 | db | ATTTTGGTGTACAGACATCACAGAATGGCCCTACCTCTCTACTTTGTTCTTACTCAAGA | 600 |
| 601 | 2Y | CCGAGACACCCCCACCTTCCCATTACTAGGGTTGAAGCCATTCGACGTGTTGAAGT | 660 |
| 601 | db | CCGAGACACCCCCACCTTCCCATTACTAGGGTTGAAGCCATTCGACGTGTTGAAGT | 660 |
| 661 | 2Y | GAAAGCAAGGGAGATTTGGTGTGCTGTGGAAGCCAGTTGCTCAATGAAATATGTGTC | 720 |
| 661 | db | GAAAGCAAGGGAGATTTGGTGTGCTGTGGAAGCCAGTTGCTCAATGAAATATGTGTC | 720 |
| 721 | 2Y | TGTCAAAATATTTCCAATACAGGACAAACAGTCTCTGCGAGAAATGAATATGAAGTCTATAG | 780 |
| 721 | db | TGTCAAAATATTTCCAATACAGGACAAACAGTCTCTGCGAGAAATGAATATGAAGTCTATAG | 780 |
| 781 | 2Y | TCTACTCGAATGAAGCATGAGAACATCTACAGTTTCAATGGTGACAGAAAAGAGCCAC | 840 |
| 781 | db | TCTACTCGAATGAAGCATGAGAACATCTACAGTTTCAATGGTGACAGAAAAGAGCCAC | 840 |
| 841 | 2Y | CAGTGTGGATGTGGACCTGTGGCTAATTCACAGCAATTTCAATGAAAAGGGCTCACTGT | 900 |
| 841 | db | CAGTGTGGATGTGGACCTGTGGCTAATTCACAGCAATTTCAATGAAAAGGGCTCACTGT | 900 |
| 901 | 2Y | CTTTCTTAAAGCTAAATGTGGTCTCTTTGGAAATCACTTTGTTCATATTTGCGAAGAACATGGC | 960 |
| 901 | db | CTTTCTTAAAGCTAAATGTGGTCTCTTTGGAAATCACTTTGTTCATATTTGCGAAGAACATGGC | 960 |
| 961 | 2Y | TAGAGGATTTGGCATATTTACATGAGGATATACCTGCTTAAAGATGGCCACAGCCTGC | 1020 |
| 961 | db | TAGAGGATTTGGCATATTTACATGAGGATATACCTGCTTAAAGATGGCCACAGCCTGC | 1020 |
| 1021 | 2Y | AATCTCTCAGGGACATCAAAAGTAAAATGTGCTTGAAGAAAACAATCTGACAGCTTG | 1080 |
| 1021 | db | AATCTCTCAGGGACATCAAAAGTAAAATGTGCTTGAAGAAAACAATCTGACAGCTTG | 1080 |
| 1081 | 2Y | CATTGTGACTTTGGTGTGGCTTAAAGTTGCGAGGCTGGCAAGTCTGCGAGTGACACCCA | 1140 |
| 1081 | db | CATTGTGACTTTGGTGTGGCTTAAAGTTGCGAGGCTGGCAAGTCTGCGAGTGACACCCA | 1140 |
| 1141 | 2Y | TGGGCAAGTTGGTACCCGGAGGTATATGGCTCCAGAGGTTGAGGGGTCTTAAACTT | 1200 |

| | | | |
|------|----|--|------|
| 61 | 2b | CTCTGGGAAAATGGGAGCTGCTGCAAAAGTTGGCGTTGCGCGTCTTCTTATCTCTTGCTC | 120 |
| 121 | 2y | TTTCAGGTGCTATACTTGGCAGATCAGAACTCAGAGTGTCTTTTCTTTAATGCTAAATG | 180 |
| 121 | 2b | TTTCAGGTGCTATCTTGGCAGATCAGAACTCAGAGTGTCTTTTCTTTAATGCTAAATG | 180 |
| 181 | 2y | GGAAARGACAGAACCAACGACTGGTGTGCAACCTTGCTATGCTGATAAAGATAAAG | 240 |
| 181 | 2b | GGAAAGACAGAACCAACGACTGGTGTGCAACCTTGCTATGCTGATAAAGATAAAG | 240 |
| 241 | 2y | GGCAGATTTGTTTGTCTACTCTGGAAGATATTTCTGGTCCATTTGAATAGTGAAGCAAG | 300 |
| 241 | 2b | GGCAGATTTGTTTGTCTACTCTGGAAGATATTTCTGGTCCATTTGAATAGTGAAGCAAG | 300 |
| 301 | 2y | TTGTTGGCTGGATGATATCAACTGCTATGACAGACTGATTTGTGTNGAAAAAAGACAG | 360 |
| 301 | 2b | TTGTTGGCTGGATGATATCAACTGCTATGACAGACTGATTTGTGTAGAAAAAAGACAG | 360 |
| 361 | 2y | CCCTGAAGTGTACTTTTGTGCTGTGAGGGCAATATGCTGTAATGAAAGTTCTCTTATTT | 420 |
| 361 | 2b | CCCTGAAGTGTACTTTTGTGCTGTGAGGGCAATATGCTGTAATGAAAGTTCTCTTATTT | 420 |
| 421 | 2y | TCCGGAGATGGAAGTCACACGCCACATTCAAATCCTGTTACACGAGGCCACCTTATTA | 480 |
| 421 | 2b | TCCGGAGATGGAAGTCACACGCCACATTCAAATCCTGTTACACGAGGCCACCTTATTA | 480 |
| 481 | 2y | CAACATTTCTGCTATTTCTTGGTACCACTAATGTTAAATTGAGGAATTTGTCATTTGTGC | 540 |
| 481 | 2b | CAACATTTCTGCTATTTCTTGGTACCACTAATGTTAAATTGAGGAATTTGTCATTTGTGC | 540 |
| 541 | 2y | ATTTGGGTGTACAGACATCACAGAATGGCCCTACCTCTCTACTTGTTCCTACTCAAGA | 600 |
| 541 | 2b | ATTTGGGTGTACAGACATCACAGAATGGCCCTACCTCTCTACTTGTTCCTACTCAAGA | 600 |
| 601 | 2y | CCGAGACACCCCCACCTTCCCATTACTAGGGTTGAAGCCATTCGACGTGTGTAAGT | 660 |
| 601 | 2b | CCGAGACACCCCCACCTTCCCATTACTAGGGTTGAAGCCATTCGACGTGTGTAAGT | 660 |
| 661 | 2y | GAAGCAAGGGGAAGATTTGGTTGTGCTGTGGAAGCCAGTTGCTCAATGATATGTGC | 720 |
| 661 | 2b | GAAGCAAGGGGAAGATTTGGTTGTGCTGTGGAAGCCAGTTGCTCAATGATATGTGC | 720 |
| 721 | 2y | TGTCAAAATATTTCCAAATACAGGACAAACGTCCTCGCAGAAATGAATATGAAGTCTATAG | 780 |
| 721 | 2b | TGTCAAAATATTTCCAAATACAGGACAAACGTCCTCGCAGAAATGAATATGAAGTCTATAG | 780 |
| 781 | 2y | TCTACCTGAAATGAAGCATGAGAACATCTACAGTTTCAATGGTGACAGAAAAAGGCGAC | 840 |
| 781 | 2b | TCTACCTGAAATGAAGCATGAGAACATCTACAGTTTCAATGGTGACAGAAAAAGGCGAC | 840 |
| 841 | 2y | CAGTGTGGATGTGGACCTGTGGCTAATCAAGCAATTTCAATGAAAAGGGCTCACTGTGAGA | 900 |
| 841 | 2b | CAGTGTGGATGTGGACCTGTGGCTAATCAAGCAATTTCAATGAAAAGGGCTCACTGTGAGA | 900 |
| 901 | 2y | CTTTCTTAAGGCTAAATGTGGTCTCTTTGGAAATCACTTTGTGTCATATTTGCGAAGACCAATGGC | 960 |
| 901 | 2b | CTTTCTTAAGGCTAAATGTGGTCTCTTTGGAAATCACTTTGTGTCATATTTGCGAAGACCAATGGC | 960 |
| 961 | 2y | TAGAGGATTTGGCATATTTACATGAGGATATACCTGCTTAAAGATGGCCACAGGCTGCT | 1020 |
| 961 | 2b | TAGAGGATTTGGCATATTTACATGAGGATATACCTGCTTAAAGATGGCCACAGGCTGCT | 1020 |
| 1021 | 2y | AATCTCTCAGGGACATCAAAAGTAAAATGCTGTTGAAAAACAATCTGACAGCTTG | 1080 |
| 1021 | 2b | AATCTCTCAGGGACATCAAAAGTAAAATGCTGTTGAAAAACAATCTGACAGCTTG | 1080 |
| 1081 | 2y | CATTGTGACTTTGGTTGGCTTAAAGTTGCGAGGCTGGCAAGTCTGCGAGTGACACCCA | 1140 |
| 1081 | 2b | CATTGTGACTTTGGTTGGCTTAAAGTTGCGAGGCTGGCAAGTCTGCGAGTGACACCCA | 1140 |
| 1141 | 2y | TGGCGAGGTTGGTACCCGGAGGTATATGGCTCCAGAGGTTGAGGGGCTCTATAACTT | 1200 |

QY 2281 TAATTACAAATTCACCTGCTGTTATTTAAGAAAAATGTAAGCTATGCTTAGTCCCAATAG 2340
 Db 2281 TAATTACAAATTCACCTGCTGTTATTTAAGAAAAATGTAAGCTATGCTTAGTCCCAATAG 2340
 QY 2341 TAAGTGGCTATTGTAAGACAGAGTGTGTTAGCTTTTCTTCTACCTGGCTTTGTAATTAGGGA 2400
 Db 2341 TAAGTGGCTATTGTAAGACAGAGTGTGTTAGCTTTTCTTCTACCTGGCTTTGTAATTAGGGA 2400
 QY 2401 AAACAAGTCTCTCTTTGAAATGGAAGAAGTAATGTTGTCACCTACCCCTACCTACCTAT 2460
 Db 2401 AAACAAGTCTCTCTTTGAAATGGAAGAAGTAATGTTGTCACCTACCCCTACCTACCTAT 2460
 QY 2461 ATCAAGGTCCTCAAAATATTTCTTTTCCATTTCAAAAGACAGACATTTGAAAACCCCTAAATTA 2520
 Db 2461 ATCAAGGTCCTCAAAATATTTCTTTTCCATTTCAAAAGACAGACATTTGAAAACCCCTAAATTA 2520
 QY 2521 CAAGCAGTAGAAGAAAGCTAAACACGCTTTTACAAATAGCC 2563
 Db 2521 CAAGCAGTAGAAGAAAGCTAAACACGCTTTTACAAATAGCC 2563

RESULT 4
 S48190 2277 bp mRNA linear ROD 08-MAY-1993
 LOCUS type II activin receptor [rats, ovary, mRNA, 2277 nt].
 DEFINITION S48190
 ACCESSION S48190.1 GI:258941
 VERSION S48190.1
 KEYWORDS Rattus sp.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.

REFERENCE 1 (bases 1 to 2277)
 AUTHORS Shinzaki, H., Ito, I., Hasegawa, Y., Nakamura, K., Igarashi, S., Nakamura, M., Miyamoto, K., Eto, Y., Ibuki, Y. and Minegishi, T.
 TITLE Cloning and sequencing of a rat type II activin receptor
 JOURNAL FEBS Lett. 312 (1), 53-56 (1992)
 MEDLINE 93050162
 PUBMED 1385212
 REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gisbq 117304] from the original journal article.
 This sequence comes from Fig. 1.

FEATURES
 source 1..2277
 /organism="Rattus sp."
 /mol_type="mRNA"
 /db_xref="taxon:10118"
 gene 1..2277
 /gene="type II activin receptor, rActR-II"
 CDS 22..1563
 /gene="type II activin receptor, rActR-II"
 /note="precursor; This sequence comes from Fig. 1;
 rActR-II"
 /codon_start=1
 /product="type II activin receptor"
 /protein_id="AA23958.1"
 /db_xref="GI:258942"
 /translation="MGAARKLAFVFLVSCSSGAILGRSETORCLFFNANWERDRTNO
 TGVEPYGDKRRHCFATKWNISIEIVKGGWLLDDINCVDRTDCEIKKDSPEYVE
 CCCEGNVCNEXFSYFPMEVETQPSNPVTPKPYNNILYSLVPLMLIAGIVICAFVW
 YRHQWVPPVLPVTPQGPFPPLGLKPLQLEVKARFGFCVWKAQLNLEYVAI
 KIFPTQDKQWNEVSVLPGMKHENILQFCAEKRGTSVDVLDLWLTAFHEKGLSL
 DFLKANVSNELCHIAETWARGLAVLHEDIPLGDKGHPAISHRDIKSNVLLKNL
 TACTADFGALKEFGKSGDTHQVGRYRMAPVELEGALNFORDAFLRDMVAMGL
 VLMELASCTAADGPVDEYMLFPFEIEIGQPSLEDQCEVHVHKKRPVLRDYQKQKAG
 MAMLCETIEECWDHDAERLACGCVGERITQWQRLNLTITTEDIVTVVTVNVDVFP
 KESSL"

ORIGIN

Query Match 81.2%; Score 2080.4; DB 10; Length 2277;
 Best Local Similarity 95.6%; Pred. No. 0;

Matches 2185; Conservative 2; Mismatches 80; Indels 19; Gaps 4;
 QY 57 GCGCGCTCGGAAATGAGAGCTGCTGAAAGTTGCGCGTTCGCGCTTTCTTATCTCTT 116
 Db 8 GCGCGCTCGGAAATGAGAGCTGCTGAAAGTTGCGCGTTCGCGCTTTCTTATCTCTT 67
 QY 117 GCTCTTCAGGTGCTATCTTTGCGAGATCAGAAACTCAGGAGTGTCTTTTCTTAAATGCTA 176
 Db 68 GCTCTTCAGGTGCTATCTTTGCGAGATCAGAAACTCAGGAGTGTCTTTTCTTAAATGCTA 127
 QY 177 ATTGGAAAAAGACAGAACCAACAGACTGGTGTGAAACCTTCTATGTTGATGAAGATA 236
 Db 128 ATTGGAAAAAGACAGAACCAACAGACTGGTGTGAAACCTTCTATGTTGATGAAGATA 187
 QY 237 AACGCGACATTTGTTTGTCTACCTGGAAGAAATTTCTTGTCTTCAATGAAATAGTGAAGC 296
 Db 188 AACGCGACATTTGTTTGTCTACCTGGAAGAAATTTCTTGTCTTCAATGAAATAGTGAAGC 247
 QY 297 AAGTTGTTGGCTGGATGATATCAACTGCTATCAGAGACTGATTTGTTGAAAAAAG 356
 Db 248 AAGTTGTTGGCTGGATGATATCAACTGCTATCAGAGACTGATTTGTTGAAAAAAG 307
 QY 357 ACAGCCCTCAAGTGTACTTTTGTGCTGAGGGCAATATGTTGTAATGAAAAAGTTCTCTT 416
 Db 308 ACAGCCCTCAAGTGTACTTTTGTGCTGAGGGCAATATGTTGTAATGAAAAAGTTCTCTT 367
 QY 417 ATTTTCCGAGATGGAAGTACACAGCCACTTCAAAATCTTGTACACCGAAGCCACCT 476
 Db 368 ATTTTCCGAGATGGAAGTACACAGCCACTTCAAAATCTTGTACACCGAAGCCACCT 427
 QY 477 ATTACAACATTCGCTGTATTCCTTGTGACCACTAAATGTTAAATGACGAAATGTCATT 536
 Db 428 ACTACAACATTCGCTGTATTCCTTGTGACCACTAAATGTTAAATGACGAAATGTCATT 487
 QY 537 GTGCATTTTGGGTGACAGACATCACAAAGATGGCTACCCCTCCTGTACTTGTTCCTACTC 596
 Db 488 GTGCGTTTGGGTGACAGACATCACAAAGATGGCTACCCCTCCTGTACTTGTTCCTACTC 547
 QY 597 AAGACCCAGACACACCCCTCCCTTCCCTTCTAGGTTGAAGCCATTGACGCTGTTAG 656
 Db 548 AAGACCCAGACACACCCCTCCCTTCCCTTCTAGGTTGAAGCCATTGACGCTGTTAG 607
 QY 657 AAGTGAAGCAAGGGGAAGATTTGGTGTGTCGGAAGACCCAGTTGCTCAATGAATAG 716
 Db 608 AAGTGAAGCAAGGGGAAGATTTGGTGTGTCGGAAGACCCAGTTGCTCAATGAATAG 667
 QY 717 TGGCTGTCAAAATATTTCCATACAGGCAACACAGTCTCTGCGAGATGAATGAAAGTCT 776
 Db 668 TGGCTATCAAAATATTTCCGATACAGGCAACACAGTCTCTGCGAGATGAATGAAAGTCT 727
 QY 777 ATAGTCTACCTGGAATGAAGCATGAGAAACATACAGTTTCAATGGTGAGAAAAAGAG 836
 Db 728 ATAGTTACCTGGATGAGCATGAAACATACATACAGTTTCAATGGTGAGAAAAAGAG 787
 QY 837 GCACAGTGTGGATGTGGACCTGTGGCTAAATCAGAGCAATTCATGAAAAGGCTCACTGT 896
 Db 788 GCACAGTGTGATGTGGACCTGTGGCTAAATCAGAGCAATTCATGAAAAGGCTCACTGT 847
 QY 897 CAGACTTTCTTAAAGCTAAATGTGGTCTCTTGAATCACTTTTGTCTATTTGCAAGAACCA 956
 Db 848 CAGACTTTCTTAAAGCTAAATGTGGTCTCTTGAATCACTTTTGTCTATTTGCAAGAACCA 907
 QY 957 TGGCTAGAGGATGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAAGC 1016
 Db 908 TGGCTAGAGGATGGCATATTTACATGAGGATATACCTGGCTTAAAGATGGCCACAAGC 967
 QY 1017 CTGCAATCTCTCACAGGACATCAAAAGTAAATAATGCTGTTTGAATAACAATCTGACAG 1076
 Db 968 CTGCAATCTCTCACAGGACATCAAAAGTAAATAATGCTGTTTGAATAACAATCTGACAG 1027
 QY 1077 CTTGCAATCTGCTGACTTTGGGTTGGCTTAAAGTTCAGGCTGCGAGTCTGCGAGTACCA 1136
 Db 1028 CTTGCAATCTGCTGACTTTGGGTTGGCTTAAAGTTCAGGCTGCGAGTCTGCGAGTACCA 1087

| | | | |
|----|------|---|------|
| 2Y | 1137 | CCNATGGCAGGTTGGTACCCGGAGGTATATGCTCCAGAGGTTGGAGGGTGCTATAA | 1196 |
| 2b | 1088 | CCCATGCCAGGTTGGTATACCCGGAGGTATATGGCTCCAGAGGTTTAGAGGGTGCTATAA | 1147 |
| 2Y | 1197 | ACTTCCAAAGGAGCGCATTTCTGAGGATAGATATGTACGCCATGGGATTAGTCTCTATGGG | 1256 |
| 2b | 1148 | ACTTCCAAAGGAGCGCATTTCTGAGGATAGATATGTACGCCATGGGATTAGTCTCTATGGG | 1207 |
| 2Y | 1257 | AATTTGGCTTCTGTTGCACTGCTGAGATGGACCCGTAGATAGTACATGTTTACCATTTG | 1316 |
| 2b | 1208 | AATTTGGCTTCTGTTGCACTGCTGAGATGGACCCGTAGATAGTACATGTTTACCATTTG | 1267 |
| 2Y | 1317 | AGGAAGAAATTTGGCCAGCATCCATCTCTTTAGAGATATGCAGGAAGTTGTTGTGCATAAA | 1376 |
| 2b | 1268 | AGGAAGAAATTTGGCCAGCATCCATCTCTTTAGAGATATGCAGGAAGTTGTTGTGCATAAA | 1327 |
| 2Y | 1377 | AAAAGAGCCCTGTTTAAAGATTTATGGCAGAAAATGTCAGGAATGGCAATGCTCTGTG | 1436 |
| 2b | 1328 | AAAAGAGCCCTGTTTAAAGATTTATGGCAGAAAATGTCAGGAATGGCAATGCTCTGTG | 1387 |
| 2Y | 1437 | AAACGATAGAAAGAAATGTTGGGATCATGATGCAAGCCAGGTTATCAGCTGGATGTGTAG | 1496 |
| 2b | 1388 | AAACGATAGAAAGAAATGTTGGGATCATGATGCAAGCCAGGTTATCAGCTGGATGTGTAG | 1447 |
| 2Y | 1497 | GTGAAAGAAATTTACTCAGATGCAAGAACTAAACAAATATCATCTACAGAGACATTTGTA | 1556 |
| 2b | 1448 | GTGAAAGAAATTTACTCAGATGCAAGAACTAAACAAATATCATCTACAGAGACATTTGTA | 1507 |
| 2Y | 1557 | CAGTGGTCACAAATGTTGCAATTTGACCTTCTCCCAAGAAATCTAGTCTATGATGGT | 1616 |
| 2b | 1508 | CAGTGGTCACAAATGTTGCAATTTGACCTTCTCCCAAGAAATCTAGTCTATGATGGT | 1567 |
| 2Y | 1617 | GGCAGCGTCTGTACACACTGAGGACCTGGGACTCTGAACTGGAGCTGCTAACTAAGGAAA | 1676 |
| 2b | 1568 | TGCACCACTGTGCAACACTGAGAACTCGGACTCTGAACTGGAGCTGCTAACTAAGGAAA | 1627 |
| 2Y | 1677 | GTGCTTAGTTGATTTTCTGTGTGAATAGTAGATGCGCTCCAGGACATGTACGCAAGCA | 1736 |
| 2b | 1628 | CTGCTTAGTTTATTTTCTGTGTGAATAGTAGAGGTGCTCCGCGACACGCTATGCAAGCA | 1687 |
| 2Y | 1737 | GGCCCTGTGGAAAGCATGGATCTGGGAGATGGATCTGGGAAACTTACTGTCATGCTCTGC | 1796 |
| 2b | 1688 | GGCCCTGTGGAAAGCATGGAT-----TGGGAGACTTCTCTGAGNGTCTGCTGC | 1733 |
| 2Y | 1797 | AGCAGAGATATGAAGAGGAGTCTAAGGAAAGAGTGCAAACTGTAAAGAACTTCTGAAAAA | 1856 |
| 2b | 1734 | AACACGGATATGAAGGGGGTCTAAGGGGAAA--CTGCGAACTGTAAAGAACTTCTGAAAAA | 1791 |
| 2Y | 1857 | TGTACTTCGAAGAAATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAACTCAAGT | 1916 |
| 2b | 1792 | CTTACAGGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAACTCAAGT | 1851 |
| 2Y | 1917 | ATTGCGAAACTGACATCAGATTTCTTAAATGTCTGTGAGAGACACTAAATTCCTTAAATG | 1976 |
| 2b | 1852 | ATTG-AAAACTGACATCAGATTTCTTAAATGTCTGTGAGAGACACTAAATTCCTTAAATG | 1910 |
| 2Y | 1977 | AACTACTGCTATTTTTTTTAAATGAAAACTTTTTCATTTTCAGATTTTAAAGGGTAACT | 2036 |
| 2b | 1911 | AACTACTGCTATTTTTTTTAAATCAAAAACCTTTTCATTTTCAGATTTTAAAGGGTAACT | 1970 |
| 2Y | 2037 | TTTTTATGCAATTTGCTGTGTTTCTTCTATAATGACTATTGTAATGCCAATGACACAGCT | 2096 |
| 2b | 1971 | TTTTTATGCAATTTGCTGTGTTTCTTATAATGACTATTGTAATGCCAATGACACAGCT | 2030 |
| 2Y | 2097 | TGTGAAATGTGTAGTGTCTGCTGTTCTGTGTACATAGTCACTAAAGTGGGGTACAGTAAA | 2156 |
| 2b | 2031 | TGTGAATGTGTAGTGTCTGCTGTTCTGTGTACATAGTCACTAAAGTGGGGTACAGTAAA | 2090 |
| 2Y | 2157 | GAGGCTTCCAAAGCAATTAATTAACCTCCCTCAAAGGTATACCTCAGTTCACCGTTGCT | 2216 |
| 2b | 2091 | GAGGCTTCCAAAGCAATTAATTAACCTCCCTCAAAGGTATACCTCAGTTCACCGTTGCT | 2150 |

| | | | |
|-----------------------|--|--|--|
| QY | 2217 | TAAATTATAAAATGAAAAACATACACAGAAATTTGAAATAAATCAGTCCATGTTTTATAACA | 2217 |
| Db | 2151 | TAAATTATAAAATGAAAAACATACACAGAAATTTGAAATAAATCAGTCCATGTTTTATAACA | 2210 |
| QY | 2277 | AGGTTAATTACAAATTCACCTGCTGTTATTTAAGAAAATGCTTAAGCTATGCTTAGTGCCA | 2336 |
| Db | 2211 | AGGT--ATTACAAACTCAGTCTGTTATTTAAGAAAAAATGCTAAGCTATGCTTAGTGCCA | 2268 |
| QY | 2337 | ATAGTA 2342 | |
| Db | 2269 | ATAGAA 2274 | |
| RESULT 5 | | | |
| LOCUS | AR116081 | 2122 bp | DNA linear PAT 16-MAY-2001 |
| DEFINITION | Sequence 4 from patent US 6132988. | | |
| ACCESSION | AR116081 | | |
| VERSION | AR116081.1 | GI:14096403 | |
| KEYWORDS | Unknown. | | |
| SOURCE | Unknown. | | |
| ORGANISM | Unclassified. | | |
| REFERENCE | 1 (bases 1 to 2122) | | |
| AUTHORS | Sugino H., Nakamura, T. and Shouji, H. | | |
| TITLE | DNA encoding a neuronal cell-specific receptor protein | | |
| JOURNAL | Patent: US 6132988-A 4 17-OCT-2000; | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..2122 | | |
| ORIGIN | /organism="unknown" | | |
| | /mol_type="unassigned DNA" | | |
| Query Match | 80.3%; | Score 2057.4; | DB 6; Length 2122; |
| Best Local Similarity | 98.6%; | Pred. No. 0; | |
| Matches 2092; | Conservative 2; | Mismatches 4; | Indels 24; Gaps 1; |
| QY | 63 | TCGGGAAAATGGGAGCTGCTGC | AAAAGTTGGCGTTCGCCGCTCTTTCTATCTCTTGCTCTT 122 |
| Db | 1 | TCGGGAAAATGGGAGCTGCTGC | AAAAGTTGGCGTTCGCCGCTCTTTCTATCTCTTGCTCTT 60 |
| QY | 123 | CAGGTGCTACTTGGCAGATCAGAA | CTCAGGAGTGTCTTTCTTTCTTAATGCTAAATGGG 182 |
| Db | 61 | CAGGTGCTACTTGGCAGATCAGAA | CTCAGGAGTGTCTTTCTTTCTTAATGCTAAATGGG 120 |
| QY | 183 | AAAAGACAGAACCAACACAGACT | GGTGTGTAACCTTGGCTATGCGTGTAAAGATAAACGGC 242 |
| Db | 121 | AAAAGACAGAACCAACACAGACT | GGTGTGTAACCTTGGCTATGCGTGTAAAGATAAACGGC 180 |
| QY | 243 | GACATTGTTTGTCTGCTGAGGAA | ATATTTCTGGTTTCCATTGAAATAGTGAACAGGTT 302 |
| Db | 181 | GACATTGTTTGTCTGCTGAGGAA | ATATTTCTGGTTTCCATTGAAATAGTGAACAGGTT 240 |
| QY | 303 | GTTGGCTGGATGATATCAACT | GCTATGACAGGACTGATGTGTGNGAAAAAAGACAGCC 362 |
| Db | 241 | GTTGGCTGGATGATATCAACT | GCTATGACAGGACTGATGTGTGNGAAAAAAGACAGCC 300 |
| QY | 363 | CTGAAGTGTACTTTTGTGCTGT | GAGGCAATATGTGTAATGAAAAAGTTCTCTTATTTTC 422 |
| Db | 301 | CTGAAGTGTACTTTTGTGCTGT | GAGGCAATATGTGTAATGAAAAAGTTCTCTTATTTTC 360 |
| QY | 423 | CGGAGATGGAAGTCACACAGCC | CACTTCAAATCTCTGTTACACCGAAGCCACCCATTACA 482 |
| Db | 361 | CGGAGATGGAAGTCACACAGCC | CACTTCAAATCTCTGTTACACCGAAGCCACCCATTACA 420 |
| QY | 483 | ACATTCTGCTGTAATCTTGGT | TACCACTAAATGTTAATTCAGGAATTTGTCATTTGTGCAT 542 |
| Db | 421 | ACATTCTGCTGTAATCTTGGT | TACCACTAAATGTTAATTCAGGAATTTGTCATTTGTGCAT 480 |
| QY | 543 | TTTGGGTGTACAGACATCACAA | GATGGGCTACCTCTCTGCTACTGTTCTTACTCAA---- 598 |
| Db | 481 | TTTGGGTGTACAGACATCACAA | GATGGGCTACCTCTCTGCTACTGTTCTTACTCAA---- 540 |

599 -----GACCCAGGACCCACCCACCTTCCCATTTACTAGGGTTGA 638
541 CTTTTATATATATAGAGGACCCAGGACCCACCCACCTTCCCATTTACTAGGGTTGA 600
639 AGCCATTGAGCTGTAGAAAGTGAAGAGGGAAGATTTGGTTGTGTCTGGAAGCC 698
601 AGCCATTGAGCTGTAGAAAGTGAAGAGGGAAGATTTGGTTGTGTCTGGAAGCC 660
699 AGTTGCTCAATGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGC 758
661 AGTTGCTCAATGAATATGTGGCTGTCAAAATATTTCCAAATACAGGACAAACAGTCTGGC 720
759 AGAATGAATATGAAGTCTATAGTCTACCTGGAATGAAGCATGAGAAATACATCTCAGTTCA 818
721 AGAATGAATATGAAGTCTATAGTCTACCTGGAATGAAGCATGAGAAATACATCTCAGTTCA 780
819 TTGGTGCAGAGAAAGAGGACACAGTGTGGATGTGGACCTGTGGCTAATACAGATTTTC 878
781 TTGGTGCAGAGAAAGAGGACACAGTGTGGATGTGGACCTGTGGCTAATACAGATTTTC 840
879 ATGAAAGGGCTCAGTGTGAGACTTTCTTAAGGCTAATGTGGTCTCTTGGATCACTTTT 938
841 ATGAAAGGGCTCAGTGTGAGACTTTCTTAAGGCTAATGTGGTCTCTTGGATCACTTTT 900
939 GTCATATTGCAGAAACCATGGCTAGAGATTTGGCATATTTACATGAGGATATACCTGGCT 998
901 GTCATATTGCAGAAACCATGGCTAGAGATTTGGCATATTTACATGAGGATATACCTGGCT 960
999 TAAAGATGGCCACAGGCTTGCATCTCTCACAGGACATCAAAAGTAAATATGTGCTGT 1058
961 TAAAGATGGCCACAGGCTTGCATCTCTCACAGGACATCAAAAGTAAATATGTGCTGT 1020
1059 TGAATCAATCTGACAGCTTGATCTGATCTTGGTGGTGGCTTAAAGTTGAGGCTG 1118
1021 TGAATCAATCTGACAGCTTGATCTGATCTTGGTGGTGGCTTAAAGTTGAGGCTG 1080
1119 GCAAGTCTGACAGTGTACACCCATGGGAGGTGTGTACCCGAGGTATATGCTCCAGAGG 1178
1081 GCAAGTCTGACAGTGTACACCCATGGGAGGTGTGTACCCGAGGTATATGCTCCAGAGG 1140
1179 TGTGGAGGGTGTATATACTTCCAAAGGACGATTTCTGAGGATAGATATGACGCCA 1238
1141 TGTGGAGGGTGTATATACTTCCAAAGGACGATTTCTGAGGATAGATATGACGCCA 1200
1239 TGGGATTTAGTCTATGGGAATTTGGCTTCTGCTGTGACCTGTGTCAGATGAGACCCGTAGATG 1298
1201 TGGGATTTAGTCTATGGGAATTTGGCTTCTGCTGTGACCTGTGTCAGATGAGACCCGTAGATG 1260
1299 AGTACATGTTACCATTTGAGGAAGAAATTTGGCCAGCATCTCTTGAAGATATGACGG 1358
1261 AGTACATGTTACCATTTGAGGAAGAAATTTGGCCAGCATCTCTTGAAGATATGACGG 1320
1359 AGTTGTTGTGCATATAAATAAAGAGGCTGTTTAAAGAGATTTATGGCAGAAACATGCAG 1418
1321 AGTTGTTGTGCATATAAATAAAGAGGCTGTTTAAAGAGATTTATGGCAGAAACATGCAG 1380
1419 GAATGGCAATGCTCTGTGAAACGATAGAAAGATTTGGGATCATGATGAGAGGACAGGT 1478
1381 GAATGGCAATGCTCTGTGAAACGATAGAAAGATTTGGGATCATGATGAGAGGACAGGT 1440
1479 TATCAGCTGGATGTAGGTGAAGAAATTTACTCAGATGAAAGACTAAACAAATATCATTA 1538
1441 TATCAGCTGGATGTAGGTGAAGAAATTTACTCAGATGAAAGACTAAACAAATATCATTA 1500
1539 CTACAGAGGACATTTGAACAGTGTCAAAATGGTGACAAATGTTGACTTTCTCTCCCAAG 1598
1501 CTACAGAGGACATTTGAACAGTGTCAAAATGGTGACAAATGTTGACTTTCTCTCCCAAG 1560
1599 AATCTACTATGATGTGGGACCCCTGTGTACACATGAGGACTGGGACTCTGAACTGGA 1658
1561 AATCTACTATGATGTGGGACCCCTGTGTACACATGAGGACTGGGACTCTGAACTGGA 1620
1659 GCTGCTAAGCTAAGGAAAGTGTCTAGTTGATTTTCTGTGTGAAATGAGTAGGATGCTCTC 1718

1621 GCTGTAAGCTAAGGAAAGTGTCTTAGTTGATTTCTGTGTGAAATGAGTAGGATGCTCTC 1680
1719 AGGACATGTACCAAGCAGCCCTTGTGGAAGCATGGATCTGGAGATGGATCTGGGAA 1778
1681 AGGACATGTACCAAGCAGCCCTTGTGGAAGCATGGATCTGGAGATGGATCTGGGAA 1740
1779 ACTTACTGCTATGCTGACGACAGATATGAAGAGAGTCTAAAGGGAAGAGCTGCAAACT 1838
1741 ACTTACTGCTATGCTGACGACAGATATGAAGAGAGTCTAAAGGGAAGAGCTGCAAACT 1800
1839 GTAAGAACTTCTGAAAATGTACTCGAAGATGTGCCCTCTCCAAATCAAGGATCTTTT 1898
1801 GTAAGAACTTCTGAAAATGTACTCGAAGATGTGCCCTCTCCAAATCAAGGATCTTTT 1860
1899 GCACCTGGCTAATCAAGTATTTGCAAACTGACATCAGATTTCTTAATGTCTGTGAGAAG 1958
1861 GCACCTGGCTAATCAAGTATTTGCAAACTGACATCAGATTTCTTAATGTCTGTGAGAAG 1920
1959 ACACAAATTCCTTAAATGAACACTGCTATTTTAAATGAAAACTTTTCAATTT 2018
1921 ACACAAATTCCTTAAATGAACACTGCTATTTTAAATGAAAACTTTTCAATTT 1980
2019 ATTTTAAAAAGGTAACCTTTTAAATGCAATTTGCTGTTGTTTCTTAATGATATTTGTA 2078
1981 ATTTTAAAAAGGTAACCTTTTAAATGCAATTTGCTGTTGTTTCTTAATGATATTTGTA 2040
2079 TGCACATACACACACTTGTGAATGTGTAGTGTGCTGCTGCTGTGTGTATACATGCTATC 2138
2041 TGCACATACACACACTTGTGAATGTGTAGTGTGCTGCTGCTGTGTGTATACATGCTATC 2100
2139 AAAGTGGGTACAGTAAAGAGG 2160
2101 AAAGTGGGTACAGTAAAGAGG 2122

RESULT 6

E15480
LOCUS E15480 2122 bp DNA linear PAT 28-JUL-1999
DEFINITION cDNA encoding activin receptor.
E15480
ACCESSION E15480.1 GI:5710163
VERSION JP 1998072497-A/3.
KEYWORDS Mus sp.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2122)
AUTHORS Sugino H, Nakamura T, and Shoji H.
TITLE RECEPTOR PROTEIN SPECIFIC TO NERVOUS CELL
JOURNAL Patent: JP 1998072497-A 3 17-MAR-1998;
TAKEDA CHEM IND LTD
COMMENT OS Mus sp. (mouse)
PN JP 1998072497-A/3
PD 17-MAR-1998
PF 25-OCT-1996 JP 1996284182
PR 27-OCT-1995 JP 95P 280939, 04-JUL-1996 JP 96P 174909 PI
SUGINO HIROSHI, NAKAMURA TAKANORI, SHOJI HIROKI PC
C07K14/705, C07H21/04, C12N15/09, C12P21/02, C12Q1/68, G01N33/53, PC
G01N33/566//
PC A61K38/00;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH 1. .2122
FT source /organism='Mus sp.'
FT /cell_line='P19'
FT 1. .8
FT 5'UTR
FT CDS
FT 9. .1574
FT /product='activin receptor'

FT 3'UTR 1575..2122.
Location/Qualifiers
1..2122
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"

FEATURES
source

Query Match 80.3%; Score 2057.4; DB 6; Length 2122;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2092; Conservative 2; Mismatches 4; Indels 24; Gaps 1;

63 TC GGGAAATGGGAGCTGCTGCAAAAGTTGGCGTTCGCGCTCTTCTTATCTCTTGTCTTT 122
1 TC GGGAAATGGGAGCTGCTGCAAAAGTTGGCGTTCGCGCTCTTCTTATCTCTTGTCTTT 60
123 CAGGTGCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAAGTAAATGGG 182
61 CAGGTGCTATCTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAAGTAAATGGG 120
183 AAAAGACAGAACCAACAGACTGTGTGTAACCTTGCTATGGTGAATGAATGAACGGC 242
121 AAAGAGACAGAACCAACAGACTGTGTGTAACCTTGCTATGGTGAATGAATGAACGGC 180
243 GACATTTGTTTGTACCTGGAAGATATTTCTGTTTCCATTTGAATAGTGAACAGGTT 302
181 GACATTTGTTTGTACCTGGAAGATATTTCTGTTTCCATTTGAATAGTGAACAGGTT 240
303 GTTGGCTGGATGATCAACTGCTATGACAGACTGATTTGTGNGAAAAAAGACAGCC 362
241 GTTGGCTGGATGATCAACTGCTATGACAGACTGATTTGTATAGAAAAAAGACAGCC 300
363 CTGAAGTGTACTTTTGTGTTGCTGAGGGCAATATGTGTAATGAAGTTCTCTTATTTTC 422
301 CTGAAGTGTACTTTTGTGTTGCTGAGGGCAATATGTGTAATGAAGTTCTCTTATTTTC 360
423 CGGAGATGGAAGTACACAGCCCACTTCAATCTCTTACCGAAGCCACCTTATACA 482
361 CGGAGATGGAAGTACACAGCCCACTTCAATCTCTTACCGAAGCCACCTTATACA 420
483 ACATTTCTGTTATCTTGTGTTGCTGACCAATATGTTAATGTCAGGAATGCTATTTGTCAT 542
421 ACATTTCTGTTATCTTGTGTTGCTGACCAATATGTTAATGTCAGGAATGCTATTTGTCAT 480
543 TTTGGGTGTACAGACATCACAGAGTGGCCCTACCTCTCTGCTACTTGTCTTACTCAA --- 598
481 TTTGGGTGTACAGACATCACAGAGTGGCCCTACCTCTCTGCTACTTGTCTTACTCAAACG 540
599 -----GACCCAGGACCAACCCCACTTCCCATTTACTAGGGTTGA 638
541 CCTTTCAATAATGATAGAGGACCCAGGACCAACCCCACTTCCCATTTACTAGGGTTGA 600
639 AGCCATTGCGCTGTAGAGTGAAGCAAGGGAAGATTTGTTGTTGCTGGAAGCCC 698
601 AGCCATTGCGCTGTAGAGTGAAGCAAGGGAAGATTTGTTGTTGCTGGAAGCCC 660
699 AGTTGCTCAATCAATATGTTGCTGCAAAATATTTCCATATACAGGACAAACAGTCTTGGC 758
661 AGTTGCTCAATCAATATGTTGCTGCAAAATATTTCCATATACAGGACAAACAGTCTTGGC 720
759 AGAATGAATATGAAGTCTATAGTCTACCTGGAATGAAGCATGAGAACTACTACAGTTCA 818
721 AGAATGAATATGAAGTCTATAGTCTACCTGGAATGAAGCATGAGAACTACTACAGTTCA 780
819 TTGTCGAGAGAAAAAGAGCAACGATGTGATGTGACCTGTGGCTAATCACAGATTTC 878
781 TTGTCGAGAGAAAAAGAGCAACGATGTGATGTGACCTGTGGCTAATCACAGATTTC 840
879 ATGAAAGGGCTCACTGTGAGATTCTTAAGGCTAATGTGGTCTCTTGGATCACTTT 938
841 ATGAAAGGGCTCACTGTGAGATTCTTAAGGCTAATGTGGTCTCTTGGATCACTTT 900
939 GTCAATTGACAGAAACCATGGCTAGAGGATTTGGCATATTTACATGAGGATATACCTGGCT 998

Db 901 GTCATATTGCAAGAACCATAGCTAGAGATTGGCATATTTATCATGAGGATATACCTGGCT 960
QY TAAAGATGGCCCAAGGCTGCAATCTCTCACAGGACATCAAAAGTAAATAATGTGCTGT 1058
Db TAAAGATGGCCCAAGGCTGCAATCTCTCACAGGACATCAAAAGTAAATAATGTGCTGT 1020
QY TGAATAAACAATCTGACAGCTTGCATTTGGGTTGGCTTTAAAGTTCGAGGCTG 1118
Db TGAATAAACAATCTGACAGCTTGCATTTGGGTTGGCTTTAAAGTTCGAGGCTG 1080
QY GCAAGTCTGAGGTGACACCCATGCGGAGGTTGGTACCCGAGGATATATGGCTCCAGG 1178
Db GCAAGTCTGAGGTGACACCCATGCGGAGGTTGGTACCCGAGGATATATGGCTCCAGG 1140
QY TGTGGAGGTTGCTATAACTTCCAAAGGACGCAATTTCTGAGGATAGATATGTAACGCA 1238
Db TGTGGAGGTTGCTATAACTTCCAAAGGACGCAATTTCTGAGGATAGATATGTAACGCA 1200
QY TGGGATTAGTCTTATGGGAATTTGGCTTCTGTTGCACTGCTGCAGATGGACCCGTAGATG 1298
Db TGGGATTAGTCTTATGGGAATTTGGCTTCTGTTGCACTGCTGCAGATGGACCCGTAGATG 1260
QY AGTACATGTTTACCATTTGAGGAAGAAATTTGGCCAGCATCCATCTCTTGAAGATATGCGG 1358
Db AGTACATGTTTACCATTTGAGGAAGAAATTTGGCCAGCATCCATCTCTTGAAGATATGCGG 1320
QY AAGTTGTTGTGCATAAAAAAGAGGCTGTTTAAAGAGATTTTGGCAGAAACATGCGAG 1418
Db AAGTTGTTGTGCATAAAAAAGAGGCTGTTTAAAGAGATTTTGGCAGAAACATGCGAG 1380
QY GAATGGCAATCTCTGTGAACGATAGAGAAATTTGGGATCATGATGCAGAACCCAGGT 1478
Db GAATGGCAATCTCTGTGAACGATAGAGAAATTTGGGATCATGATGCAGAACCCAGGT 1440
QY TATCAGCTGGATGTCTAGGTGAAAGAAATTTACTCAGATGCAAGAACTAAACAATATCATTA 1538
Db TATCAGCTGGATGTCTAGGTGAAAGAAATTTACTCAGATGCAAGAACTAAACAATATCATTA 1500
QY CTACAGAGCAATTTGTAACAGTGGTCAATGGTGAACAATGTTGACTTTCTCTCCCAAG 1598
Db CTACAGAGCAATTTGTAACAGTGGTCAATGGTGAACAATGTTGACTTTCTCTCCCAAG 1560
QY AATCTAGTCTATGATGGTGGCAACCGTCTGTACACACTGAGGACTGGAAGTCTGCTCC 1658
Db AATCTAGTCTATGATGGTGGCAACCGTCTGTACACACTGAGGACTGGAAGTCTGCTCC 1620
QY GCTGCTAAGCTAAGGAAGTGTCTTAGTTGATTTTCTGTGAAATGATGAGTGGCTCTCC 1718
Db GCTGCTAAGCTAAGGAAGTGTCTTAGTTGATTTTCTGTGAAATGATGAGTGGCTCTCC 1680
QY AGGACATGTACGCAAGCAGCCCTTGTGGAAGCATGGATCTGGAGATGGATCTGGGAA 1778
Db AGGACATGTACGCAAGCAGCCCTTGTGGAAGCATGGATCTGGAGATGGATCTGGGAA 1740
QY ACTTACTGCATCGTCTGCAGCAGATATGAAGAGAGTCTAAGGGGAAAGCTGCAAACT 1838
Db ACTTACTGCATCGTCTGCAGCAGATATGAAGAGAGTCTAAGGGGAAAGCTGCAAACT 1800
QY GTAAGAACTTCTGAAAATGTACTCGAAGAAATGTGGCCCTCTCCAAATCAAGGATCTTTT 1898
Db GTAAGAACTTCTGAAAATGTACTCGAAGAAATGTGGCCCTCTCCAAATCAAGGATCTTTT 1860
QY GGACCTGGCTAATCAAGTATTTGCAAACTGACATCAGATTTCTTAATGTCTGTGAGAG 1958
Db GGACCTGGCTAATCAAGTATTTGCAAACTGACATCAGATTTCTTAATGTCTGTGAGAG 1920
QY ACATAATTCCTTAAATGAATGAACTACTGCTATTTTTTTTTTAAATGAAAACTTTTCATTTT 2018
Db ACATAATTCCTTAAATGAATGAACTACTGCTATTTTTTTTTTAAATGAAAACTTTTCATTTT 1980
QY ATTTTAAAGAGGTAACCTTTTATTTGCTATTTGCTGTTTCTTATAATGACTATTTGTA 2078

JOURNAL Biochem. Biophys. Res. Commun. 184 (1), 310-316 (1992)
 MEDLINE 92231944
 PUBMED 1314589
 COMMENT Original
 FEATURES source
 source text: Homo sapiens male testis cDNA to mRNA.
 Location/Qualifiers
 1..2382
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="male"
 /tissue_type="testis"
 174..1715
 /codon_start=1
 /product="activin type II receptor"
 /protein_id="AAA35504.1"
 /db_xref="GI:178050"
 /translation="WGAAAKLAFVFLISCSGAILGRSETQECILFFNANWEKDRNQF
 TGVEFCYGDKRRHCFCATKNIISGIEIVKQWLLDINCYDRDTCVEKQDSPEYF
 CRCHKNWKEKPSVPEVDEVTQSNPTPKPPYNNILYSLVPLMLIAGIVICAFW
 YRHGMKAVPVLPVTPDPPSPPLGLKPLQLLEVKARGFRGCVWKAQLLNEYAV
 KIPFIQDKSQWNEYEVSLPKMKHENILQIFGAEKRTGSDVDLMLITAFHEKGSLS
 DFLANFVWNEISCHIAETMARGLAYLHEDIPLGXDKHPAISHRDIKSNVLLKNNL
 TACTADFGIALKFEAGKAGDTGQVGRYMAPEVLEGAINFQDAFIRIDMYANNL
 VLWELASCTAADGPVDEYMLPFEIEIGHPSELDQMEVNVHKKKPSVLDRDYWKHAG
 MAMLCETIEECWDHDAEARLSAGCVGERITQMRLTNIITTEDIVTVVTVWTVNVDPPF
 KESSL"
 174..230
 231..1712
 /product="activin type II receptor"
 Query Match 74.2%; Score 1903; DB 9; Length 2382;
 Best Local Similarity 92.0%; Pred. No. 0;
 Matches 2089; Conservative 2; Mismatches 153; Indels 26; Gaps 7;
 1 CTCGAGGAGAACCCGAGGAACCTGGATATCTACGAGAACTTCTAGCGCTTCTCGGCG 60
 104 CTCGAGGAGAACCCGAGGAACCTGGATATCTACGAGAACTTCTCGGCTTCTCGGCG 163
 61 CCTCGGGAAGTGGGAGCTGTCRAAGTGGCGTTCGCGCTCTTCTATCTCTGCTC 120
 164 CCTCGGGAAGTGGGAGCTGTCRAAGTGGCGTTCGCGCTCTTCTATCTCTGCTC 223
 121 TTCAGTGTCTACTTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTG 180
 224 TTCAGTGTCTACTTTGGTAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTG 283
 181 GGAAARGACAGAACCAACAGACTGGTGTGAACTTGTCTATGGTGATAAAGTAAACG 240
 284 GGAAARGACAGAACCAATCAAACTGGTGTGAAACCGTGTATGGTGACAAAGTAAACG 343
 241 GGCAATCTTTTGTCTACCTGGGAAGTAATTTCTGGTTCATTGTAATAGTGAAGCAAGG 300
 344 GCGGCATTGTTTGTCTACCTGGGAAGTAATTTCTGGTTCATTGTAATAGTGAAGCAAGG 403
 301 TTGTTGGCTGGATGATATCACTGCTATGACAGGACTGATGTGTGTAAGAAAAAGACAG 360
 404 TTGTTGGCTGGATGATATCAACTGCTATGACAGGACTGATGTGTGTAAGAAAAAGACAG 463
 361 CCTGGAAGTGTACTTTTGTGCTGTGAGGCAATATGTTGAATGAAAGTCTCTTATTT 420
 464 CCTGGAAGTGTATATTTTGTGCTGTGAGGCAATATGTTGAATGAAAGTCTCTTATTT 523
 421 TCGGAGATGGAAGTGCACAGCCCACTTCAAACTCCCTGTTACACCGAAGCCACCCTATTA 480
 524 TCGGAGATGGAAGTGCACAGCCCACTTCAAACTCCGATTTACCTAAGCCACCCTATTA 583
 481 CAACATCTGCTGTATTTCTTGGTGGTACCACTAATGTTAATTTGCAGGAATTTGCTTGTGC 540
 584 CAACATCTGCTGTATTTCTTGGTGGTACCACTAATGTTAATTTGCAGGAATTTGCTTGTGC 643
 541 ATTTGGGTGCAGACATCACAAAGATGGCCCTACCCCTCGTGTACTTGTTCCTACTCAAGA 600

1681 TT--AGTTGATTTCTGTGTAAGTGTAGTGGTCTCCAGACATGTACGCAAGCAGC 1738
 1784 TTACAGTTTATTTCTGTGTAAATGAGTAGATGCTCTTGGAAATGTTAAGAAAG 1843
 1739 CCTTGTGGAAGCATGATCTGGAGATGGATCTGGAAACTTACTGATCTCTGCAG 1798
 1844 ACCCTTTGTGAAA-----ATGTTGCTCTGGAGACTTACTGCAATTCGACAG 1893
 1799 CACAGATATGAG---AGGAGTCTAGGAAAGCTGCAACCTGTAAGAACTTCTG-- 1852
 1894 CACAGATGGAAGGACATGAGACTAAGAAACCTTGCAAACTCTATAAGAAACTTTG 1953
 1853 --AAATGTACTCGAAGATGTGGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAA 1910
 1954 AAAAGTGTACATGAGATGTAGCCCTCTCCAAATCAAGGATCTTTTGGACCTGGCTAA 2013
 1911 TCAAGTATTTGCAAACTGACATGATCTTTTAAATGCTGTGTGAGAGACATTAATCTCT 1970
 2014 TGGAGTGTGTTG-AAACTGACATGATCTTTTAAATGCTGTGTGAGAGACATTAATCTCT 2072
 1971 TAAATGAACTACTGCTATTTTAAATGCAAAACTTTTCAATTCAGATTTTAAAGG 2030
 2073 TAAATGAACTACTGCTATTTTAAATGCAAAACTTTTCAATTCAGATTTTAAAGG 2132
 2031 GTAAC---TTTTATTGCAATTTCTGTTGTTCTTATAAATGACTATTTGTAATGCCAAT 2087
 2133 GTAACTTGTTTATTGCAATTTCTGTTGTTCTTATAAATGACTATTTGTAATGCCAAT 2192
 2088 GACACAGCTTGTGAATCTGTAGTGTCTGCTGTTCTCTGTACAT--AGTCATCAAGTGG 2145
 2193 GACACAGCTTGTGAATCTGTAGTGTCTGCTGTTCTCTGTACAT--AGTCATCAAGTGG 2252
 2146 GTTACAGTAAAGAGGTTCCAGCATTAATTAACCTCCCTCAACAGGTTATACCTCAGT 2205
 2253 GTTACAGTAAAGAGGTTCCAGCATTAATTAACCTCCCTCAACAGGTTATACCTCAGT 2312
 2206 TCCACGGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2255
 2313 TCCACGGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2362

RESULT 10
 HUMACTRIIA 2375 bp mRNA linear PRI 07-FEB-2003
 LOCUS
 DEFINITION Homo sapiens ActRIIA mRNA for activin typeII A receptor, complete cds.
 ACCESSION D31770
 VERSION D31770.1 GI:1321631
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Donaldson,C.J., Mathews,L.S. and Vale,W.W.
 TITLE Molecular cloning and binding properties of the human type II activin receptor
 JOURNAL Biochem. Biophys. Res. Commun. 184 (1), 310-316 (1992)
 MEDLINE 92231944
 PUBMED 1314589
 REFERENCE 2 (bases 1 to 2375)
 AUTHORS Iimura,T. and Oida,S.
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 2375)
 AUTHORS Iimura,T.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUN-1994) Tadashi Iimura, Faculty of Dent., Tokyo Medical and Dental University, Department of Biochemistry; 1-5-45 Yushima, Bunkyo-ku, Tokyo 113, Japan (Tel:81-3-3813-6111(ex.5126), Fax:81-3-5684-8047)
 COMMENT On May 16, 1996 this sequence version replaced gi:577623.
 FEATURES Location/Qualifiers

1. 2375
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="ACTR2MG"
 /cell_line="MG-63"
 /cell_type="osteosarcoma"
 /clone_lib="lambda gt11"
 1. 2375
 /gene="ActRIIA"
 1. 82
 /gene="ActRIIA"
 /note="site of divergence"
 164. 1705
 /gene="ActRIIA"
 /codon_start=1
 /product="activin typeII A receptor precursor"
 /protein_id="BAA06548.1"
 /db_xref="GI:1321632"
 /translation="MGAAKLAFAVFISSGAILGRSEOCLEFNANWEKDRTNQ
 TGVEPCYGDKKRHCFAWKNGISIEIVKQCNLDDINCYDRDCEVKKDSBEVVF
 CCEGNCMEKESYFPEMEVTPQTSNPVTPKPPYNNILLYSLVPLMLIAGIVICAFWV
 VRHKMAYPVLPVTPQDPGPPPLGLKPLQLEVKARGPSPSLKQAQLNEXYAV
 KTFTQDKSQWNEVEVSLPGMKHENILQIFGAEKGTSDVDLWITTAFFHEKGSLS
 DFKANYVSNELCHIAETWARGLAIVLHEDIPLGKQKPAISHRDIKSNVLLKNL
 TACIADFGUALKFAVAGSAGDTGQVGTTRYMAPEVLEGAINFORDAFLRDMYAMGL
 VLWELASRCTAADGPVDEYMLPFEETIGQHPSEDMQEVVVKKRPVLRDYWKHAG
 MAMLCETIEECWDHDAEARLSAGCVGERITQMQLTNIITTEDIVTVVMTNVDPPP
 KESSL"
 164. 220
 /gene="ActRIIA"
 221. 1702
 /gene="ActRIIA"
 /product="activin typeII A receptor"
 2351. 2356
 /gene="ActRIIA"
 2375
 /gene="ActRIIA"
 sig_peptide 73.88; Score 1892.4; DB 9; Length 2375;
 mat_peptide 91.58; Pred.No.0;
 polyA_signal 2; Mismatches 164; Indels 29; Gaps 8;
 polyA_site 2375
 ORIGIN
 Query Match 73.88; Score 1892.4; DB 9; Length 2375;
 Best Local Similarity 91.58; Pred.No.0;
 Matches 2098; Conservative 2; Mismatches 164; Indels 29; Gaps 8;
 QY 1 CTCGGAGGAAGACCCAGGGAACCTCGATATCTAGCGAGAACTTCCTACGCGTCTTCGGCGG 60
 Db 94 CTCGGAGGAAGACCCAGGGAACCTCGATATCTAGCGAGAACTTCCTACGCGTCTTCGGCGG 153
 QY 61 CCTCGGAAAAATGGAGCTGCTGCAAAAGTTGGCGTTCGGCGTCTTCCTATCTCTGCTC 120
 Db 154 CCTCGGAAAAATGGAGCTGCTGCAAAAGTTGGCGTTCGGCGTCTTCCTATCTCTGCTC 213
 QY 121 TTCAGGTGCTATCTTGGCAGATCAGAAACTCAGGAGTGTCTTTCTTTTAAATGCTAATTG 180
 Db 214 TTCAGGTGCTATCTTGGTATAGTATCAGAAACTCAGGAGTGTCTTTCTTTTAAATGCTAATTG 273
 QY 181 GGAAAAAGCAGACAGACCAACAGACTGGGTGTTGAACCTTGTCTATGTTGATAAAGATAAAG 240
 Db 274 GGAAAAAGCAGACAGACCAACAGACTGGGTGTTGAACCTTGTCTATGTTGATAAAGATAAAG 333
 QY 241 GCAGCATTTGTTTGTCTACCTCGAAGAAATATTTCTGTTCCATTTGAATAGTGAAGCAAG 300
 Db 334 GCAGCATTTGTTTGTCTACCTCGAAGAAATATTTCTGTTCCATTTGAATAGTGAAGCAAG 393
 QY 301 TTGTTGGCTGATGATATCAATCTGATGACAGGACTGATGTTGTTGTAAGAAAAAGACAG 360
 Db 394 TTGTTGGCTGATGATATCAATCTGATGACAGGACTGATGTTGTTGTAAGAAAAAGACAG 453
 QY 361 CCTTGAAGTGTACTTTTGTGTTGAGGCAATATGTTGTAAGAAAAAGTCTCTTATTT 420
 Db 454 CCTTGAAGTGTACTTTTGTGTTGAGGCAATATGTTGTAAGAAAAAGTCTCTTATTT 513
 QY 421 TCCGAGATGGAAGTACACAGACCCACTTCAATCTCTGTTTACACCGAAGCCACTTATA 480

514 TCCGGAGATGGAAGTCAACAGCCCACTTCAATCCAGTTACACCTTAAGCCACCCCTATTA 573
481 CAACATCTGCTGATCTCTTCTGTTACCACTAATGTTAATTCAGGAAATGTCATTTGTGC 540
574 CAACATCTGCTGATCTCTTCTGTTACCACTAATGTTAATTCAGGAAATGTCATTTGTGC 633
541 ATTTTGGGTGTACAGACATCAAGATGCGCTTACCTCTCTGTTACTTGTTCCTACTCAAGA 600
634 ATTTTGGGTGTACAGGATCAAGATGCGCTTACCTCTCTGTTACTTGTTCCTACTCAAGA 693
601 CCCAGGACCCACCCACCTTCCCACTTACTAGGTTGAAGCCATTCAGCTGTTAGAGT 660
694 CCCAGGACCCACCCACCTTCCCACTTACTAGGTTGAAGCCATTCAGCTGTTAGAGT 753
661 GAAAGCAGGGGAGATTTGTTGTTGTTCTGTTGAAAGCCAGTTGCTCAATGAATATGTGGC 720
754 GAAAGCAGGGGAGATTTGTTGTTGTTCTCTGTTGAAAGCCAGTTGCTCAATGAATATGTGGC 813
721 TGTCAAAATATTTCCAAATACAGGACAAACAGTCTCTGTCAGATGAATATGAAGTCTATAG 780
814 TGTCAAAATATTTCCAAATACAGGACAAACAGTCTCTGTCAGATGAATATGAAGTCTATAG 873
781 TCTACCTGGAATGAAGCATGAGAACATACAGTTTCATTTGTTGTCAGAGAAAGAGGCAC 840
874 TTTGCTGGAATGAAGCATGAGAACATATACAGTTTCATTTGTTGTCAGAGAAAGAGGCAC 933
841 CAGTGTGATGTGACCTGTGCTAATCAAGCATTTCAATGAAGGCTCAGTGTACAGA 900
934 CAGTGTGATGTGATCTTTGCTGATCAAGCATTTCAATGAAGGCTCAGTGTACAGA 993
901 CTTTCTTAAGGTAATGTTGTTCTTGGATCACTTTGTCATTTGTCAGAAACCATGTC 960
994 CTTTCTTAAGGTAATGTTGTTCTTGGATCACTTTGTCATTTGTCAGAAACCATGTC 1053
961 TAGAGATTTGGCATATTTACATGAGGATATACCTGCTTAAAGATGGCCACAGCCTGC 1020
1054 TAGAGATTTGGCATATTTACATGAGGATATACCTGCTTAAAGATGGCCACAGCCTGC 1113
1021 AATCTCTCACAGGACATCAAAAGTAAATGTTGCTTTGTTGAAAGAAATCTGACAGCTTG 1080
1114 CATATCTCACAGGACATCAAAAGTAAATGTTGCTTTGTTGAAAGAAATCTGACAGCTTG 1173
1081 CATCTGCTGATTTGGTTGGCTTAAAGTTGAGGCTGGCAAGTCTGAGGTGACACCCA 1140
1174 CATCTGCTGATTTGGTTGGCTTAAAGTTGTTGCTGGCAAGTCTGAGGTGACACCCA 1233
1141 TGGCAGGTGTTGATCCCGAGGTATATGGCTCCAGAGGTCTTGGAGGTGCTATAAATTT 1200
1234 TGGCAGGTGTTGATCCCGAGGTATATGGCTCCAGAGGTATATGGAGGTGCTATAAATTT 1293
1201 CCAAGGGAGCGATTTCTGAGGATAGATATGTCGCCATGGGATTTAGTCTTATGGGAAT 1260
1294 CCAAGGGAGCGATTTCTGAGGATAGATATGTCGCCATGGGATTTAGTCTTATGGGAAT 1353
1261 GGCTTCTCCTGTCACCTGTCAGATGGACCCGTAGATGAGTACATGTTACCATTTGAGGA 1320
1354 GGCTTCTCCTGTCACCTGTCAGATGGACCCGTAGATGAGTACATGTTACCATTTGAGGA 1413
1321 AGAATTTGGCAGCATCCATCTCTTGAAGATATGACAGGAGTTGTTGTCATATAAAAAA 1380
1414 GGAATTTGGCAGCATCCATCTCTTGAAGATATGACAGGAGTTGTTGTCATATAAAAAA 1473
1381 GAGGCTGTTTGAAGATATTTGGCAGAACATGTCAGGATGCAATGCTCTGTGAAAC 1440
1474 GAGGCTGTTTGAAGATATTTGGCAGAACATGTCAGGATGCAATGCTCTGTGAAAC 1533
1441 GATAGAGATTTGGATCATGATGACAGAGCCAGGTTATCAGCTGGATGTTAGGTGA 1500
1534 CATTGAAGATTTGGATCAGCAGCGAGAGCCAGGTTATCAGCTGGATGTTAGGTGA 1593
1501 AAGAATTACTCAGATGCAAGACCTTAAACATATCATTTACTACAGAGGACATTTGAACAGT 1560

1594 AAGAAATTACCAGATGCAGAGACTAAACAAATATTATTACCACAGAGGACATTTGTAACAGT 1653
1561 GGTCAATGTTGTGACAAATGTTGACTTTCCTCCCAAGAAATCTAGTCTATGATGTTGGCA 1620
1654 GGTCAATGTTGTGACAAATGTTGACTTTCCTCCCAAGAAATCTAGTCTATGATGTTGGC 1713
1621 CCCTCTGTACACACTGAGGACTGGGACTCTGAAGTGGAGCTGCTAAGCTAAGGAAGTGC 1680
1714 CCATGTGTGCACACTAAGAAATGGGACTCTGAACCTGGAGCTGCTAAGCTAAGAAACTGC 1773
1681 TT--AGTTGATTTTCTGTGTAATGAGTAGGATGCTCCAGGACATGTACGCAAGCAGC 1738
1774 TTACAGTTTATTTCTGTGTAATGAGTAGGATGCTCTTGGAAATGTTAAGAAAGAAG 1833
1739 CCCTTGTGGAAGCAGTGGATCTGGAGATGAGTCTGGGAAAATTACTGCACTGCTGAG 1798
1834 ACCCTTCTGTGAAA-----ATGTTGCTCTGGGAGACTTACTGCACTTGCCTGAG 1883
1799 CACAGATATGAAG---AGGAGTCTAAGGGAAGAGCTGCAAACTGTA-----AAGAACTTC 1850
1884 CACAGATGTGAAGGACATGAGACTAAGAGAAACCTTTGCAAACTCTATAAGAAAACTTTT 1943
1851 TGAAATGCTACTCGAAGATGCGCCCTCTCCAAATCAAGGATCTTTTGGACCTTGGGTAA 1910
1944 AAAAGTGTACATGAAGAAATGAGCCCTCTCCAAATCAGAGATCTTTTGGACCTGGCTAA 2003
1911 TCAAGTATTTGCAAACTGACATCAGATTTCTTAATGCTCTGTGTCAGAGACATTAATTCCT 1970
2004 TGGAGTGTG--AAAACCTGACATCAGATTTCTTAATGCTGTGTCAGAGACATTAATTCCT 2062
1971 TAAATGAACTACTGCTATTTTAAATGAAAAAATTCTTCAATTCAGATTTTAAAGG 2030
2063 TAAATGAACTACTGCTATTTTAAATGAAAAAATTCTTCAATTCAGATTTTAAAGG 2122
2031 GTAAC---TTTTATGCAATTTGCTGTTGTTCTTATAAATGACTATTGTAATGCCAAT 2087
2123 GTAACCTGTTTATGCAATTTGCTGTTGTTCTTATAAATGACTATTGTAATGCCAAT 2182
2088 GACACAGCTTGTGAATGCTGTAGTGTGCTGTGTTCTGTGTACAT--AGTCAATCAAAGTGG 2145
2183 GACACAGCTTGTGAATGTTTAGTGTGCTGTGTTCTGTGTACATTAAGTCAATCAAAGTGG 2242
2146 GGTACAGTAAAGAGGCTTCCAGCATTTACTTTAACTCCCTCCCAAGGATATCCTCAGT 2205
2243 GGTACAGTAAAGAGGCTTCCAGCATTTACTTTAACTCCCTCCCAAGGATATCCTCAGT 2302
2206 TCCACGGTTGTTAAATATAAATGAAACATTAACAGATTTG---AATAATCAGTTC 2262
2303 TCCACGGTTGTTAAATATAAATGAAACATTAACAGATTTGAAATATAATCGATC 2362
2263 CATGTTTATAAC 2275
2363 CATGTTTGTAAAC 2375

RESULT 11
ARL16086
LOCUS
DEFINITION
SEQUENCE 11 from patent US 6132988.
ARL16086
ACCESSION
ARL16086.1 GI:14096408
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 (bases 1 to 1563)
Sugino, H., Nakamura, T. and Shouji, H.
DNA encoding a neuronal cell-specific receptor protein
Patent: US 6132988-A 11 17-OCT-2000;
Location/Qualifiers
1. .1563
/organism="unknown"
/mol_type="unassigned DNA"

ARL16086
Sequence 11 from patent US 6132988.
ARL16086
ARL16086.1 GI:14096408
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 1563)
Sugino, H., Nakamura, T. and Shouji, H.
DNA encoding a neuronal cell-specific receptor protein
Patent: US 6132988-A 11 17-OCT-2000;
Location/Qualifiers
1. .1563
/organism="unknown"
/mol_type="unassigned DNA"

[illegible]

sig peptide

ORIGIN

| | | | | |
|----------------------------|--------|----------------|-----------|--------------|
| Query Match | 57.2%; | Score 1466; | DB 10; | Length 1543; |
| Best Local Similarity | 96.8%; | Pred. No. 0; | | |
| Matches 1493; Conservative | 2; | Mismatches 48; | Indels 0; | Gaps 0; |

| | | | |
|-----|-----|--|-----|
| 67 | GA | AAATGGGAGCTGCTGCAAAAGTTGGGGTTCGGCGTCTTTCTTATCTCTCTGCTCTTCAGG | 126 |
| 1 | GA | AAATGGGAGCTGCTGCAAAAGTTGGCGTTTCGGCGTCTTTCTTATCTCTCTCTTCAGG | 60 |
| 127 | TG | CTATACCTGGCAGATCAGAACTCAGGAGTGCTTTTCTTAACTGCTAAATGGGAAA | 186 |
| 61 | TG | CTATACCTGGCAGATCGGAACTCAGGAGTGCTTTCTTAACTGCTAAATGGGAAG | 120 |
| 187 | RG | CAGAAACCAACACAGACTGGTGTGAACCTTGCTATGCTGATAAAGATAAACGGCGACA | 246 |
| 121 | AG | CAGAACCAATCAGACTGGTGTGAGCCTTGCTATGGTGATAAAGCAAAACGACGACA | 180 |
| 247 | TT | GTGTTTGCTACCTGGAAGATAATTTCTGGTTCATTTGAATAAGTGAAGCAAGTTGTG | 306 |
| 181 | TT | GTGTTTGCTACCTGGAAGATAATTTCTGGTTCATTGAAATAGTAAGCAAGTTGTG | 240 |
| 307 | GC | TGGATGATATCAACTGCTATGACAGGACTGATTGTGTGTAAGAAAAAGACAGCCCTGA | 366 |
| 241 | GC | TGGATGATATCAACTGCTATGACAGGACTGATTGTATGAAAAAAGACAGCCCTGA | 300 |
| 367 | AG | TGACTTTTGTGCTGTGAGGGCANATGTGTAATGAAAGTCTCTTATTTTCCGGA | 426 |
| 301 | AG | TGACTTTTGTGCTGTGAGGGCANATGTGTAATGAAAGTCTCTTATTTTCCGGA | 360 |
| 427 | GAT | GGAAGTCACACAGCCCACTTCAAATCTGTGTTACACCGAAGCCACCCTATTACAACAT | 486 |
| 361 | GAT | GGAAGTCACACAGCCCACTTCAAATCTGTGTTACCCGGAACCAACCTACTACACAT | 420 |
| 487 | TC | TGCTGATATTCCTTGGTACCACTAAATGTTAAATTGCAGGAATGTCTATTTGCAATTTG | 546 |
| 421 | CT | TGCTGATATTCCTTGGTACCACTATGTTAAATTGGGGGATTGTCAATTTGTCGTTTTG | 480 |
| 547 | GG | TGTACAGACATCACAAAGATGGCTACCCCTCTGTACTTGTTCCTACTCAAGACCCAGG | 606 |
| 481 | GG | TGTACAGACATCACATGATGGCTACCCCTCCTGTACTTGTTCCTACTCAGGACCCCTGG | 540 |
| 607 | AC | ACCCCCACCTTCCCGATTTACTAGGGTTGAAGCCATTGACGCTGTTAGAAGTGAAGAC | 666 |
| 541 | ACC | ACCTTCCACCTTCCCGATTTACTAGGGTTGAAGCCATTGACGCTTATTGGAAGTGAAGAC | 600 |
| 567 | AG | GGGAGATTTGGTTGTGCTCGGAAGCCAGTTGCTCAATGAATATGTGGCTGTCTAA | 726 |
| 601 | AG | GGGAGATTTGGTTGTGCTCGGAAGCCAGTTGCTCAATGAATATGTGGCTGTCTAA | 660 |
| 727 | AA | TATTTCCCAATACAGGACAAACAGTCCCTGGCAGAAATGAATATGAAGCTATAGTCTACC | 786 |
| 661 | AA | TATTTCCGATACAGGACAAACAGTCCCTGGCAGAAATGAATATGAAGCTATAGTCTACC | 720 |
| 787 | TG | GAATGAAGCATGAGAACATACTACAGTTTCAATTTGGTCAGAGAAAAAGGACCAGTGT | 846 |

[illegible]

RESULT 13
BD015780 1442 bp DNA linear
LOCUS BD015780
psFINITION Novel protein and utilization thereof. PAT 27-AUG-2002

| SOURCE | MUS sp. |
|-----------|---|
| ORGANISM | Mus sp. |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| AUTHORS | 1 (bases 1 to 1442) |
| TITLE | Sugino, H. |
| JOURNAL | Novel protein and utilization thereof |
| COMMENT | Patent: JP 2001204478-A 1 31-JUL-2001; TAKEDA CHEMICAL INDUSTRIES LTD OS Mus sp. (mouse) PN JP 2001204478-A/1 |

PD 31-JUL-2001
PF 27-OCT-2000 JP 2000333544
PI HIROSHI SUGINO
PC C12N15/09,A61K31/711,A61K38/00,A61K45/00,A61P25/08,A61P25/14,
PC A61P25/16,
PC A61P25/28,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,G01N33/15,G01N33/50,G01N33/566//A01K67/027,C12P21/08, PC
C1201/68,
PC C12P21/02,C12P1:645,C12N15/00,A61K37/02,C12N5/00 CC Novel
protein and utilization thereof
FH Key Location/Qualifiers
FT source 1. 1442
FT /organism='Mus sp. (mouse)'.
FEATURES
source
1. 1442
/organism='Mus sp.'
/mol_type='genomic DNA'
/db_xref='taxon:10095'
ORIGIN
Query Match 56.2%; Score 1440; DB 6; Length 1442;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1440; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 554 AGACATCACAGATGGCTACCCCTCTGTACTTGTCTCTACTCAAGACCCAGGACACCC 613
DB 1 AGACATCACAGATGGCTACCCCTCTGTACTTGTCTCTACTCAAGACCCAGGACACCC 60
QY 614 CCACCTCCCATCTACTAGGGTTGAAGCCATTGAGCTGTAGAGTGAAGCAAGGGGA 673
DB 61 CCACCTCCCATCTACTAGGGTTGAAGCCATTGAGCTGTAGAGTGAAGCAAGGGGA 120
QY 674 AGATTGGTGTCTGGAAGCCAGTTGCTCAATGAATATGTGGCTGTCAAAATATTT 733
DB 121 AGATTGGTGTCTGGAAGCCAGTTGCTCAATGAATATGTGGCTGTCAAAATATTT 180
QY 734 CCAATCAGGACAAACAGTCTCGGAGAAATGAATGAAGTCTAGTCTACTCGGAATG 793
DB 181 CCAATCAGGACAAACAGTCTCGGAGAAATGAATGAAGTCTAGTCTACTCGGAATG 240
QY 794 AAGCATGAGAACATCTACAGTCTTCAATGAAGGCTCACTCTCAGACTTCTTAAGGCT 853
DB 241 AAGCATGAGAACATCTACAGTCTTCAATGAAGGCTCACTCTCAGACTTCTTAAGGCT 300
QY 854 GACCTGGCTTAATCAGACAGATTTTCATGAAAGGCTCACTCTCAGACTTCTTAAGGCT 913
DB 301 GACCTGGCTTAATCAGACAGATTTTCATGAAAGGCTCACTCTCAGACTTCTTAAGGCT 360
QY 914 AATGTGCTCTCTTGGAAATCARCTTTGTTCATATTTGAGAAACCATGGCTAGAGGATGGCA 973
DB 361 AATGTGCTCTCTTGGAAATCARCTTTGTTCATATTTGAGAAACCATGGCTAGAGGATGGCA 420
QY 974 TATTTACATGAGATATACCTGGCTTAAAGATGAGCAAGCCATGCAATCTCTCAGG 1033
DB 421 TATTTACATGAGATATACCTGGCTTAAAGATGAGCAAGCCATGCAATCTCTCAGG 480
QY 1034 GACATCAAAAGTAAATGTGCTGTGTAAGCAATCTCAGAGTTGCAATGCTGACTTT 1093
DB 481 GACATCAAAAGTAAATGTGCTGTGTAAGCAATCTCAGAGTTGCAATGCTGACTTT 540
QY 1094 GGGTGGCTTAAAGTTCGAGGCTGCGAGTCTGAGGTTGACACCCATGGCGAGGTTGGT 1153
DB 541 GGGTGGCTTAAAGTTCGAGGCTGCGAGTCTGAGGTTGACACCCATGGCGAGGTTGGT 600
QY 1154 ACCCGAGGTATATGGCTCCAGAGTGTGGAGGTTGCTATAACTTCCAAAGGGACGCA 1213
DB 601 ACCCGAGGTATATGGCTCCAGAGTGTGGAGGTTGCTATAACTTCCAAAGGGACGCA 660
QY 1214 TTTCTGAGATAGATATGATACCCCATGGAGTATGCTCTATGGAAATGGCTTCTCGTTGC 1273

DB 661 TTTCTGAGGATAGATATGTAGCCATGGGATTAGTCTTATGGGAATGGCTTCTCGTTGC 720
QY 1274 ACTGCTGCGAGATGAGCCCGTAGATGAGTACATGTTTACATTTAGGAAGAAATTTGCCAG 1333
DB 721 ACTGCTGCGAGATGAGCCCGTAGATGAGTACATGTTTACATTTAGGAAGAAATTTGCCAG 780
QY 1334 CATCCATCTCTTGAAGATATGAGGAAAGTTGTTGTGATATAAAAAAGAGGCTGTTTTA 1393
DB 781 CATCCATCTCTTGAAGATATGAGGAAAGTTGTTGTGATATAAAAAAGAGGCTGTTTTA 840
QY 1394 AGAGATATTTGGCAGAAACATGAGGAATGGAATGCTCTGTAACCATGAGAAATGTT 1453
DB 841 AGAGATATTTGGCAGAAACATGAGGAATGGAATGCTCTGTAACCATGAGAAATGTT 900
QY 1454 TGGGATCATGATGAGAAAGCCAGGTTATCAGCTGGATGTGTAGTGAAGAAATTTACTCAG 1513
DB 901 TGGGATCATGATGAGAAAGCCAGGTTATCAGCTGGATGTGTAGTGAAGAAATTTACTCAG 960
QY 1514 ATGCAAGACTAAACAATATCATTTACTACAGAGACATTTAAACAGTGGTCAATGGTG 1573
DB 961 ATGCAAGACTAAACAATATCATTTACTACAGAGACATTTAAACAGTGGTCAATGGTG 1020
QY 1574 ACAATGTTGACTTCTCTCCAAAGAAATCTAGTCTATGATGGTGGCACCGCTGTACACA 1633
DB 1021 ACAATGTTGACTTCTCTCCAAAGAAATCTAGTCTATGATGGTGGCACCGCTGTACACA 1080
QY 1634 CTGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAAGTCTTAGTTGATTTTC 1693
DB 1081 CTGAGGACTGGGACTCTGAACTGGAGCTGCTAAGCTAAGGAAAGTCTTAGTTGATTTTC 1140
QY 1694 TGTGTGAATGAGTGGATGCTCCAGGACATGTACGCAAGCAGCCCTTTGGAAAGCA 1753
DB 1141 TGTGTGAATGAGTGGATGCTCCAGGACATGTACGCAAGCAGCCCTTTGGAAAGCA 1200
QY 1754 TGGATCTGGGAGATGAGTCTGGGAAACTTACTGCAATGCTGTCAGCACAGATATGAGAG 1813
DB 1201 TGGATCTGGGAGATGAGTCTGGGAAACTTACTGCAATGCTGTCAGCACAGATATGAGAG 1260
QY 1814 GAGTCTAAGGAAAGCTGCAAACTGTAAAGAACTTCTGAAATCTACTCGAGAGATGTG 1873
DB 1261 GAGTCTAAGGAAAGCTGCAAACTGTAAAGAACTTCTGAAATCTACTCGAGAGATGTG 1320
QY 1874 GCCTCTCCCAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTTGCAAACTGACAT 1933
DB 1321 GCCTCTCCCAATCAAGGATCTTTTGGACCTGGCTTAATCAAGTATTTGCAAACTGACAT 1380
QY 1934 CAGATTTCTTAATGCTGTGCAAGACACTAATTCCTTAAATGAACTACTGCTATTTTTT 1993
DB 1381 CAGATTTCTTAATGCTGTGCAAGACACTAATTCCTTAAATGAACTACTGCTATTTTTT 1440
QY 1994 TT 1995
DB 1441 TT 1442
RESULT 14
BD095083
LOCUS Novel protein and its use.
DEFINITION Novel protein and its use.
ACCESSION BD095083
VERSION BD095083.1 GI:22640671
KEYWORDS WO 0131004-A/1.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Sugino,H.
REFERENCE
1 (bases 1 to 1442)
AUTHORS
TITLE
Novel protein and its use
PATENT: WO 0131004-A 1 03-MAY-2001;
JOURNAL
TAKEDA CHEMICAL INDUSTRIES LTD,HIROSHI SUGINO
COMMENT
OS Mus sp. (mouse)
PN WO 0131004-A/1

| | |
|----------------|---|
| PD | 03-MAY-2001 |
| Pf | 26-OCT-2000 WC 2000JP007480 |
| PR | 27-OCT-1999 JP 99P 305187, 04-NOV-1999 JP 99P 313453 PI |
| Hiroshi Sugino | |
| PC | C12N15/12,C12N5/10,C12P21/02,C07K16/18,C07K14/47,G01N33/50, PC GOIN33/15, |
| CC | A61K38/17 |
| CC | Novel protein and its use |
| FH | Key Location/Qualifiers |
| FT | source 1..1442 /organism='Mus sp. (mouse)' |
| FEATURES | Location/Qualifiers |
| source | 1..1442 /organism="Mus sp." /mol_type="genomic DNA" /db_xref="taxon:10095" |

ORIGIN

| Query Match | | 56.2% | Score 1440; | DB 6; | Length 1442; |
|----------------------------|------|---|--------------|-------|-------------------|
| Best Local Similarity | | 99.9% | Pred. No. 0; | | |
| Matches 1440; Conservative | | 1; | Mismatches | 1; | Indels 0; Gaps 0; |
| QY | 554 | AGACATCACAGATGGCCTACCTCCTGTACTTGTTCCTACTCAAGACCCAGACGACACCC | 613 | | |
| DB | 1 | AGACATCACAGATGGCCTACCTCCTGTACTTGTTCCTACTCAAGACCCAGACGACACCC | 60 | | |
| QY | 614 | CCACCTTCCCATTACTAGGGTTGAAGCCATTGCAGCTGTTAGAAGTGAAGCAAGGGGA | 673 | | |
| DB | 61 | CCACCTTCCCATTACTAGGGTTGAAGCCATTGCAGCTGTTAGAAGTGAAGCAAGGGGA | 120 | | |
| QY | 674 | AGATTTGGTTGTCTGGAAAGCCAGTGTCTCAATGAATATATGGTGTCAAAATATTT | 733 | | |
| DB | 121 | AGATTTGGTTGTCTGGAAAGCCAGTGTCTCAATGAATATATGGTGTCAAAATATTT | 180 | | |
| QY | 734 | CCAATACAGGACAAACAGTCCTCGCGAATGAATATGAAGTCTATAGCTACCTGGAATG | 793 | | |
| DB | 181 | CCAATACAGGACAAACAGTCCTCGCGAATGAATATGAAGTCTATAGCTACCTGGAATG | 240 | | |
| QY | 794 | AAGCATGAGAACATATCTACAGTTTCAATGGTGCAGAGAAAGACCCAGTGTGGATGTG | 853 | | |
| DB | 241 | AAGCATGAGAACATATCTACAGTTTCAATGGTGCAGAGAAAGACCCAGTGTGGATGTG | 300 | | |
| QY | 854 | GACCTGTGGCTAATCAACAGATTTCAATGAAGGGCTCACTGTCAAGACTTCTTAAGGCT | 913 | | |
| DB | 301 | GACCTGTGGCTAATCAACAGATTTCAATGAAGGGCTCACTGTCAAGACTTCTTAAGGCT | 360 | | |
| QY | 914 | AATGTGCTCTTTGGGAATCAACCTTTGTTCATATTGCAGAAACCATGGCTAGAGGATGGCA | 973 | | |
| DB | 361 | AATGTGCTCTTTGGGAATCAACCTTTGTTCATATTGCAGAAACCATGGCTAGAGGATGGCA | 420 | | |
| QY | 974 | TATTTACATGAGANATATCCTGGCTTAAAGATGGCCACAAGCCTGCAATCTCTCACAGG | 1033 | | |
| DB | 421 | TATTTACATGAGANATATCCTGGCTTAAAGATGGCCACAAGCCTGCAATCTCTCACAGG | 480 | | |
| QY | 1034 | GACATCAAAAGTAAATATGTCTTGTGAAATAAATCTGCAGAGCTTCCATTTGTGACTTT | 1093 | | |
| DB | 481 | GACATCAAAAGTAAATATGTCTTGTGAAATAAATCTGCAGAGCTTCCATTTGTGACTTT | 540 | | |
| QY | 1094 | GGGTGTGCCCTTAAAGTTTCGAGGCTGGCAAGTCTCTCAGGTGACACCCATGGCGAGTTGGT | 1153 | | |
| DB | 541 | GGGTGTGCCCTTAAAGTTTCGAGGCTGGCAAGTCTCTCAGGTGACACCCATGGCGAGTTGGT | 600 | | |
| QY | 1154 | ACCGGAGGTATATGGCTCCAGAGGTTGTTGGAGGGTCTATAAATCTTCCAAAGGGACGCA | 1213 | | |
| DB | 601 | ACCGGAGGTATATGGCTCCAGAGGTTGTTGGAGGGTCTATAAATCTTCCAAAGGGACGCA | 660 | | |
| QY | 1214 | TTTCTGAGGATAGATATGTACGCCATGGGATTTAGTCCCTATGGGAATTTGGCTTCTCGTTGC | 1273 | | |
| DB | 661 | TTTCTGAGGATAGATATGTACGCCATGGGATTTAGTCCCTATGGGAATTTGGCTTCTCGTTGC | 720 | | |
| QY | 1274 | ACTGCTGCAGATGGACCCGTAGATAGATATATGTTACATTTTGGGAAGAAATTTGCCAG | 1333 | | |
| DB | 721 | ACTGCTGCAGATGGACCCGTAGATAGATATATGTTACATTTTGGGAAGAAATTTGCCAG | 780 | | |

| | | | |
|------|----|--|------|
| 974 | QY | TATTTACATGAGGATATACCTGGCTTTAAAAAGATGGCCACAAGCCTGCAATCTCTCACAGG | 1033 |
| 421 | Db | TATTTACATGAGGATATACCTGGCTTTAAAAAGATGGCCACAAGCCTGCAATCTCTCACAGG | 480 |
| 1034 | QY | GACATCAAAGTAAAAATGTGCTGTTGAAAAAACAATCTGCAGGTTGCATTTGCTGACTTT | 1093 |
| 481 | Db | GACATCAAAGTAAAAATGTGCTGTTGAAAAAACAATCTGCAGGTTGCATTTGCTGACTTT | 540 |
| 1094 | QY | GGCTTGGCCCTTAAAGTTCGAGGGCTGGCAACTCTGCAGGTGACACCCCATGGCAGGTTGGT | 1153 |
| 541 | Db | GGGTGGCCCTTAAAGTTCGAGGGCTGGCAAGTCTGCAGGTGACACCCCATGGCAGGTTGGT | 600 |
| 1154 | QY | ACCCGGAGGTATATGGCTCCAGAGGTCTTGGAGGGTGCTATAAACTTCCAAAGGGACGCA | 1213 |
| 601 | Db | ACCCGGAGGTATATGGCTCCAGAGGTCTTGGAGGGTGCTATAAACTTCCAAAGGGACGCA | 660 |
| 1214 | QY | TTTCTGAGGATAGATATGTACGCCATGGGATTAGTCCTATGGGAATTGGCTTCTCGTTGC | 1273 |
| 661 | Db | TTTCTGAGGATAGATATGTACGCCATGGGATTAGTCCTATGGGAATTGGCTTCTCGTTGC | 720 |
| 1274 | QY | ACTGCTGCAGATGGACCCGTAGATAGATACATGTTACATTTGAGGAAGAAATTGGCCAG | 1333 |
| 721 | Db | ACTGCTGCAGATGGACCCGTAGATAGATACATGTTACATTTGAGGAAGAAATTGGCCAG | 780 |

TITLE Direct Submission
 JOURNAL Submitted (13-DEC-1995) Nora Flavin, Pharmacology, Biotechnology
 Centre, University College Dublin, Belfield, Dublin, 4, Ireland
 FEATURES
 source Location/Qualifiers
 1..1670
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /chromosome="2"
 /map="2q2.3-2.4"
 /tissue_type="ovary"
 CDS
 1..1542
 /function="serine/threonine kinase"
 /codon_start=1
 /product="activin receptor type IIA"
 /protein_id="AAC48694.1"
 /db_xref="GI:1809244"
 /translation="MGAAAKLAFAPVLISCSGAILGRSETQECIFYNANWERDRTNR
 TCVECYGKDKRRXCFKATKNISGIEIVKQGWLDINDCYDRDTCLEKQSPFVY
 CGCEGNMNERSTYFPEMEVTOPTGNPTPKPPYNNILLISLVLMCLAGIVICAFVY
 YRHKKWAPVPLVPTQDPGPPSPPLGLKPLQLLEVKARGRGFCWKQAQLLNEVAV
 KLPFQDKQSWNEVEVSLPKMKHENILQFIGAEKRGTSVDVDLWLTAFPEKSGSL
 DIKFNQVSNWELCHTAFMARGLAYLHIEDPLKDGHPKAPISHRDIKSNVLLKNNL
 TACIDFGLALFKFAGSKSDGTGVOGTRYMAPEVLGAINFOADFRLIDMYAMGL
 WIELASRGTADGPDVEYMLPFEIEIGQHSLEDMQEVVHKRPFLVRIDYWKHAG
 MAMLCETIECDHDAEARLSAGCVERITQMQLNTIITTEDIVTVMTVNTVDFPP
 KESSL"
 ORIGIN
 Query Match 56.1%; Score 1436.6; DB 4; Length 1670;
 Best Local Similarity 91.5%; Pred. No. 0;
 Matches 1530; Conservative 2; Mismatches 137; Indels 3; Gaps 1;
 QY 71 ATGGGAGCTGCTGCAAAAGTTGGCGTTGGCGCTTTCTTATCTCTTCTTCAGGTGCT 130
 Db 1 ATGGGAGCTGCTGCAAAAGTTGGCGTTGGCGCTTTCTTATCTCTTCTTCAGGTGCT 60
 QY 131 ATACTTTGGCAGATCAGAACTCAGGAGTGTCTTTCTTTAATGCTAATTTGGGNAARGAC 190
 Db 61 ATACTTTGGTAGATCAGAACTCAGGAGTGTATTTTCTATAATGCTAATTTGGGNAAGAT 120
 QY 191 AGAACCAACACAGACTCGTGTTGAACCTTGCTATGGTGATAAAGATAAAGCGGCACATTGT 250
 Db 121 AGAACGAATCGCACTGGTGTGTGAATCCTGTTCGGTGAACAAGATAACCGCGGCATTTGT 180
 QY 251 TTTTGCTACCTGGGAAGAATATTTCTGGTTCCATTGAAATAGTGAAGCAAGGTTCTTGGCTG 310
 Db 181 TTTTGCTACGTGGGAAGAATATTTCTGGTTCCATTGAAATAGTGAAGCAAGGTTCTTGGCTG 240
 QY 311 GATGATATCAACTGCTATGACAGGACTGATGTGTNGAAAAAAGACAGCCCTGAAGTG 370
 Db 241 GATGATATCAACTGCTATGACAGGACTGATGTGTATAGAAAAAAGATAGCCCTGAAGTA 300
 QY 371 TACTTTTGTGCTGTGAGGCAATATGTGTAAATGAAAAGTTCTCTTATTTTCGGAGATG 430
 Db 301 TATTTTGTGCTGTGAGGCAATATGTGTAAACGAAAGGTTTTCTTATTTTCGGAGATG 360
 QY 431 GAAGTCACACAGCCCACTTCAATCTCTGTTTACACCGAAGCCACCCATTACACATTTCTG 490
 Db 361 GAGGTACACAGCCCTACGTGAATPCAGTAACCTTAAAGCCTCCCTATTACAAACATCTTG 420
 QY 491 CTGTATTCTCTTGGTACCACTAATGTGTAAATTCGAGGAATTTGCAATTTTGGGTTG 550
 Db 421 CTCTATTCTTTGGTGCCACTTATGTTAATTTGGGGGATTTGTAATTTGTGCATTTTGGGTTG 480
 QY 551 TACAGACATCACAGATAGGCGTCACTCCCTCTGTACTTGTTCCTACTCAAGACCCAGGACCA 610
 Db 481 TACAGACATCATAGATGGCTTACCTCTCTGTGCTCGTTCCTCAACTCAAGACCCAGGACCA 540
 QY 611 CCCCACCTTCCCATTACTTAGGGTTGAAGCCATTGACGCTGTTAGAGTGAAGCAAGG 670
 Db 541 CCCCACCTTCTCCATTACTAGGTTTGAAGCCATTACAGTTACTAGAGTGAAGCAAGG 600

Search completed: February 26, 2004, 05:45:04
Job time : 9695 secs